

Titre: Nouvelle approche pour le dimensionnement et l'optimisation de la gestion des réservoirs et de la qualité de l'eau potable dans les réseaux de distribution
Title:

Auteur: Essoyeke Batchabani
Author:

Date: 2014

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Batchabani, E. (2014). Nouvelle approche pour le dimensionnement et l'optimisation de la gestion des réservoirs et de la qualité de l'eau potable dans les réseaux de distribution [Ph.D. thesis, École Polytechnique de Montréal]. PolyPublie. <https://publications.polymtl.ca/1380/>
Citation:

 **Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/1380/>
PolyPublie URL:

Directeurs de recherche: Musandji Fuamba
Advisors:

Programme: Génie civil
Program:

UNIVERSITÉ DE MONTRÉAL

NOUVELLE APPROCHE POUR LE DIMENSIONNEMENT ET
L'OPTIMISATION DE LA GESTION DES RÉSERVOIRS ET DE LA QUALITÉ
DE L'EAU POTABLE DANS LES RÉSEAUX DE DISTRIBUTION

ESSOYEKE BATCHABANI

DÉPARTEMENT DES GÉNIES CIVIL, GÉOLOGIQUE ET DES MINES

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION

DU DIPLÔME DE PHILOSOPHIÆ DOCTOR

(GÉNIE CIVIL)

AVRIL 2014

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée :

NOUVELLE APPROCHE POUR LE DIMENSIONNEMENT ET
L'OPTIMISATION DE LA GESTION DES RÉSERVOIRS ET DE LA QUALITÉ
DE L'EAU POTABLE DANS LES RÉSEAUX DE DISTRIBUTION

présentée par : BATCHABANI Essoyeke

en vue de l'obtention du diplôme de : Philosophiæ Doctor

a été dûment acceptée par le jury d'examen constitué de :

M. MAHDI Tew-Fik, Ph.D., président

M. FUAMBA Musandji, Ph.D., membre et directeur de recherche

M. BOUAZZA Zoubir, Ph.D., membre

M. FILION Yves R., Ph.D., membre

DÉDICACE

Je dédie ce travail, aboutissement d'un long processus, à toutes les entités et personnes contributrices parmi lesquelles :

Mon Papa : *qui ne cesse d'intercéder pour nous*

Ma maman : *qui nous a donné l'éducation de base*

Mes frères et sœurs: *avec qui on a toujours partagé les projets*

Toute ma famille : *qui m'a toujours encouragé et soutenu*

Mes collègues : *qui m'ont toujours facilité l'intégration*

Mes professeurs : *qui m'ont transmis du savoir*

Mon Directeur du Collège : *qui a su encourager mon excellence jusqu'au bout*

Mon pays : *qui m'a donné la vitesse initiale*

Le Maroc : *qui m'a accueilli et assuré ma formation d'ingénieur*

Le Canada : *qui m'a accueilli et assuré mes recherches*

Moi-même : *pour mon courage*

REMERCIEMENTS

Avant d'aller plus loin, j'aimerais mettre à profit la présente occasion pour remercier toutes les personnes qui m'ont aidées et soutenues tout au long de ce parcours qui tire à sa fin.

Je souhaite remercier particulièrement mon directeur de recherche, M. Musandji Fuamba, pour son apport inconditionnel sur tous les plans tout au long du processus ayant mené au stade final de la thèse. Ses qualités de confiance, d'orientation, d'écoute et d'encouragement ont été plus que favorables pour la réussite de ce projet. En dehors même du présent projet, il m'a formé sur plusieurs plans et m'a donné de précieux conseils.

Mes remerciements s'adressent ensuite à Nicolas Basile dont le mémoire de maîtrise a servi de point de départ pour la présente thèse. Sa contribution a permis de cerner certains aspects de la problématique et les défis à relever dans ce domaine complexe du génie.

Je tiens à remercier tous les membres du jury qui ont accepté d'évaluer ce travail. Leurs disponibilités et leurs suggestions sont énormément appréciées. Je remercie le professeur Tew-Fik Mahdi pour tous ses précieux conseils. Je tiens également à remercier le professeur Yves R. Filion qui a accepté faire le déplacement depuis Kingston pour participer à l'évaluation de cette thèse, malgré son emploi du temps très chargé.

Un grand merci à tous mes collègues et ami(e)s qui ont contribué de près ou de loin au bon déroulement et à la finalisation de ce travail. Je remercie particulièrement Samba Bousso qui a facilité mon accueil et mon intégration au Canada. Je remercie également Nasrin Javadi pour son encouragement et nos précieux échanges sur tous les plans. Aussi, j'aimerais remercier mes collègues suivants pour leurs commentaires, suggestions, collaboration ou soutien : Abdou Seck, Nicolas Géhéniau, Fadoua Houssa, Guillaume Lamothe, Mathieu Roy, Julia Ayombo, Anas Sebti, Najib Mhamdi, France-Line Dionne, Jean-Pierre Aka, Wilfried Koblan, Paul Zang, Paul Akakpo et Clémence Djangbedja. J'exprime également ma reconnaissance à M. Yves Gosselin et Mme. Maryam Kamalinezhad pour leur présence et pour tous les bons moments qu'on a passés ensemble dans le cours d'Hydrologie.

Je souhaite remercier ma famille spirituelle ici au Canada. Un grand merci à James Poon pour la correction de l'anglais dans mes articles. Merci à toute mon équipe de prière : Ann Lee, Tu Lan Diep, Hubert Woo, Karen Wong, Ben Leung et Zoey Gow.

J'aimerais remercier mes professeurs de l'École Hassania des Travaux Publics qui m'ont permis de faire connaissance avec mon directeur de recherche. Je remercie particulièrement les professeurs Saïd Rhouzlane et Zoubir Bouazza qui ont facilité mon admission à l'École Polytechnique de Montréal. Je remercie aussi Ilham Ennaouri qui m'a aidé dans les formalités administratives.

Pour terminer cette liste qui est loin d'être exhaustive, je transmets mes vifs remerciements à toute ma famille pour son amour indéfectible. Si les mots me manquent pour exprimer ma proximité et ma reconnaissance, mon cœur ne peut se taire, et s'il arrivait même que mon cœur se taise, alors mon silence prendrait la relève.

RÉSUMÉ

Les Réseaux de Distribution d'Eau Potable (RDEP) sont traditionnellement conçus sans intégrer les réservoirs dans le processus d'optimisation, malgré les nombreux avantages que ces réservoirs apportent à l'exploitation des réseaux. Ces avantages incluent la gestion efficace des demandes, la continuité de la distribution d'eau en cas de défaillance du système, la lutte contre les incendies ou la modulation des débits de pompage. Lorsque bien conçus et bien placés dans les systèmes de distribution, les réservoirs permettent d'améliorer les performances globales de ces systèmes. Cependant, la conception des réservoirs est complexe du fait que plusieurs facteurs sont à considérer avec plusieurs variables mises en jeu.

L'approche traditionnelle utilisée pour la conception des réservoirs dans les réseaux de distribution consiste à opérer un choix manuel de l'emplacement et l'élévation pour le réservoir, suivi de l'optimisation des conduites, le volume du réservoir étant manuellement déterminé après une simulation de longue durée ou avec les guides techniques. Cette approche permet seulement un nombre limité d'essais sur l'emplacement du réservoir, son élévation ou son volume avant le choix. Elle ne permet donc pas d'explorer adéquatement l'espace des solutions et il devient très difficile d'utiliser cette approche pour les grands réseaux nécessitant beaucoup d'efforts avant de trouver une solution faisable. La tendance générale dans cette approche est de recourir à des réservoirs de grandes dimensions supposés hydrauliquement sécuritaires pour les réseaux mais qui peuvent se révéler problématiques en ce qui concerne la qualité de l'eau distribuée.

Ce projet de recherche a pour but le développement d'un modèle d'optimisation multiobjective des réseaux de distribution d'eau potable capable de prédire la nécessité d'introduction ou non d'un ou de plusieurs réservoirs dans le réseau ainsi que les emplacements de ces éventuels réservoirs.

Pour atteindre l'objectif de la recherche, la première étape a consisté en une analyse critique structurée des différentes approches incluant la considération des réservoirs dans l'optimisation des réseaux. Cette analyse a permis d'une part de confirmer la nécessité d'une prise en compte des réservoirs dans le processus d'optimisation des réseaux et d'autre part de proposer un protocole afin de permettre l'identification de la meilleure approche ou du type de modèle approprié en fonction d'un certain nombre de critères prédéterminés.

Dans un souci de proposer un modèle dont le champ d'application est large, la deuxième étape a consisté en une proposition d'un certain nombre de paramètres de références pouvant permettre la généralisation de l'approche de conception. Ces paramètres sont : (1) le nombre de zones de pression; (2) les élévations minimales et maximales des réservoirs (3) le diamètre de référence pour les conduites; (4) le débit et la hauteur manométrique de référence pour les pompes et (5) la sélection des paramètres de l'algorithme génétique. Ils sont utiles pour les applications pratiques de l'algorithme d'optimisation aux situations réelles.

Afin de fournir un outil d'aide à la décision aux gestionnaires de réseaux, la troisième étape a été consacrée au développement du modèle numérique lui-même. Une version de ce modèle est implémentée en environnement déterministe et l'autre en stochastique pour permettre une comparaison de deux approches de conception et en même temps pour donner plus de choix aux utilisateurs. La qualité de l'eau a été considérée comme faisant partie des objectifs de la conception optimale des réseaux dans les deux approches.

Le modèle développé a été appliqué à trois études de cas afin de le valider. Les résultats ont montré que la prise en compte de la qualité de l'eau permet une intégration efficace des réservoirs dans le processus d'optimisation. Bien que n'étant pas la plus réaliste, l'approche déterministe propose des solutions qui peuvent être obtenues en beaucoup moins de temps comparée à l'approche stochastique qui peut s'avérer non applicable sur des réseaux très grands et complexes.

Mots clés : réseau de distribution, eau potable, écoulement sous pression, réservoir, localisation du réservoir, simulation, dimensionnement, optimisation, optimal, multiobjective, prise de décision, algorithme génétique, incertitude, stochastique, qualité de l'eau, Latin Hypercube, corrélation.

ABSTRACT

Water Distribution Networks (WDN) are traditionally designed without integrating tanks in the optimization process, despite the numerous benefits tanks bring to network operation. These benefits include efficient demand management, water supply in case of system failure, fire fighting or the modulation of pump flow rate. When properly designed and located, storage tanks are a cost-effective means of improving overall network performance. However, tank design is complex because several factors are to be considered with several variables involved.

The traditional approach to design tanks in distribution systems is to introduce the tank into a WDN at a predetermined location and elevation (manual selection), followed by pipe optimization, the tank volume being decided through extended-period simulation (manual check) or by guidelines. This approach allows for only a limited number of trials on tank location, elevation and storage prior to selection. It does not adequately explore the solution space and it becomes very difficult to use this approach for large networks requiring much effort before finding a feasible solution. This approach tends to implement large water tanks which are supposed hydraulically safe for water distribution systems. Those large tanks can however be a source of trouble regarding the quality of the water provided to the citizens.

This research project aims to develop a model for multi-objective optimization of water distribution systems that is able to predict the need for the introduction or not of one or more tanks in the network and the locations of these potential tanks.

In order to achieve the aim of the research, the first step consisted of a review of different approaches to including tank consideration in WDN optimization. The resulting analysis first confirms the need for tank consideration in the network design optimization process and secondly led to suggest a procedure for selecting the most suitable approach to real-world WDN optimization projects based on a number of predetermined criteria.

For the sake of proposing a model with a wide application scope, the second step consisted of providing a certain number of reference parameters that may allow the design approach generalization. These parameters are: (1) the number of pressure zones; (2) the minimum and maximum tank bottom elevations; (3) the reference diameter for pipes; (4) the reference flow

and head for pumps and (5) selection of genetic algorithm parameters. They are useful for practical applications of the optimization algorithm to real practices.

In order to provide a decision support tool to network managers, the third stage was devoted to the development of the numerical model itself. A version of this model is implemented in a deterministic environment and the other in a stochastic one to allow for a comparison of both design approaches and also to give more choices to the users. The water quality was considered as part of the objectives in the network optimal design problem in both approaches.

The model was applied to three case studies in order to validate it. The results showed that including water quality in the objectives allows for an effective integration of the tanks in the optimization process. Although not being the most realistic, the deterministic approach provides solutions that can be obtained in much less time compared to the stochastic approach which may be non feasible on very large and complex networks.

Keywords : water distribution network, drinking water, pressurized flow, water tank or reservoir, tank location, simulation, design, optimization, optimal, multi-objective, decision-making, genetic algorithm, uncertainty, stochastic, water quality, Latin Hypercube, correlation.

TABLE DES MATIÈRES

DÉDICACE.....	III
REMERCIEMENTS	IV
RÉSUMÉ.....	VI
ABSTRACT	VIII
TABLE DES MATIÈRES	X
LISTE DES TABLEAUX.....	XVII
LISTE DES FIGURES	XIX
LISTE DES SIGLES ET ABREVIATIONS	XXI
LISTE DES ANNEXES	XXIII
CHAPITRE 1 INTRODUCTION.....	1
1.1 Contexte de l'étude.....	1
1.2 Optimisation de la conception des réseaux de distribution d'eau potable	1
1.3 Motivation de la recherche	3
1.4 Hypothèses scientifiques de la recherche.....	4
1.5 Objectifs de la recherche	5
1.6 Plan de la thèse	6
1.7 Justification des méthodes et paramètres utilisés dans la thèse	6
CHAPITRE 2 ARTICLE 1: OPTIMAL TANK DESIGN IN WATER DISTRIBUTION NETWORKS: REVIEW OF LITERATURE AND PERSPECTIVES	8
2.1 Présentation de l'article	8
2.1.1 Objectif.....	8
2.1.2 Conclusion.....	8

2.2	Optimal tank design in water distribution networks: review of literature and perspectives	10
2.2.1	Abstract	10
2.2.2	Introduction	10
2.2.3	Current practice to tank design in WDN	12
2.2.3.1	Type	12
2.2.3.2	Location	12
2.2.3.3	Levels	13
2.2.3.4	Volume	13
2.2.4	Optimization methods application	17
2.2.4.1	Single-Objective Models	18
2.2.4.2	Multiobjective Models	20
2.2.4.3	Resolution methods	23
2.2.5	Evaluation of both practices	24
2.2.5.1	Closing the gap between engineering practice and mathematical optimization models	24
2.2.5.2	Strengths and weaknesses of approaches to tank design in WDN	26
2.2.6	Choosing the appropriate model	28
2.2.7	Discussion	31
2.2.8	Conclusion and recommendations	31
CHAPITRE 3	ARTICLE 2: MULTI-OBJECTIVE DESIGN AND DECISION-MAKING PROCESS IN WATER DISTRIBUTION SYSTEMS	32
3.1	Présentation de l'article	32
3.1.1	Objectif	32
3.1.2	Conclusion	33

3.2	Multi-objective design and decision-making process in water distribution systems	34
3.2.1	Abstract	34
3.2.2	Introduction	34
3.2.3	Approach	36
3.2.3.1	Identify and formulate the design options and bounds	37
3.2.3.2	Define and formulate the design performance evaluation criteria	41
3.2.3.3	Solve the multi-objective optimization problem	43
3.2.4	Case study	49
3.2.4.1	Network description	49
3.2.4.2	Application results and discussions.....	52
3.2.5	Conclusion.....	60
CHAPITRE 4 ARTICLE 3: MULTI-OBJECTIVE STOCHASTIC DESIGN OF WATER DISTRIBUTION SYSTEMS INCLUDING WATER QUALITY.....		61
4.1	Présentation de l'article	61
4.1.1	Objectif.....	61
4.1.2	Conclusion.....	62
4.2	Multi-objective stochastic design of water distribution systems including water quality	63
4.2.1	Abstract	63
4.2.2	Introduction	64
4.2.3	Methodology	67
4.2.3.1	Identify and formulate the decision variables and set bounds	67
4.2.3.2	Define and formulate the design performance evaluation criteria	71
4.2.3.3	Solve the multi-objective stochastic optimization problem	76
4.2.4	Case study	81

4.2.4.1	Network description	81
4.2.4.2	Application results and discussions.....	84
4.2.5	Conclusion.....	90
CHAPITRE 5	MODÉLISATION DU RÉSEAU	91
5.1	Introduction	91
5.2	Conceptualisation du réseau de distribution.....	91
5.2.1	Objets physiques	91
5.2.1.1	Objet Nœud	92
5.2.1.2	Objet Arc	92
5.2.1.3	Objet Point de passage	93
5.2.1.4	Objet Texte.....	93
5.2.2	Objets non physiques	93
5.2.2.1	Objet Demande.....	93
5.2.2.2	Objet Courbe de modulation	93
5.2.2.3	Objet Courbe	94
5.2.2.4	Objet Commande simple	94
5.2.2.5	Objet Commande élaborée	94
5.2.2.6	Objet Énergie.....	94
5.2.2.7	Objet Source	95
5.2.2.8	Objet Réaction.....	95
5.2.2.9	Objet Temps	95
5.2.2.10	Objet Rapport	95
5.2.2.11	Objet Options	95
5.2.2.12	Objet Fond d'écran.....	95

5.3	Présentation générale du modèle	95
5.3.1	Préparation des données par l'utilisateur	96
5.3.2	Lecture et enregistrement des données par le programme	97
5.3.3	Traitement préliminaire des données	98
5.3.3.1	Calcul des zones de pression.....	98
5.3.3.2	Calcul du diamètre de référence.....	99
5.3.4	Simulation EPANET et enregistrement des résultats.....	100
5.3.5	Traitement des résultats et évaluation de la solution.....	100
5.3.6	Modélisation des variables du problème	100
5.3.7	Codage et interprétation des chaines de solution	101
5.3.7.1	Classes de base de la bibliothèque GALib (Genetic Algorithm Library)	101
5.3.7.2	Implémentation de la bibliothèque GALib au projet de recherche	104
5.3.8	Liste de quelques fonctions et leurs rôles dans le modèle.....	104
5.3.8.1	Échantillonnage avec indépendance.....	104
5.3.8.2	Échantillonnage avec corrélation	105
5.3.8.3	Conversion Diamètre Économique	105
5.3.8.4	Redondance	105
5.3.8.5	Résilience du réseau	106
5.3.8.6	Temps de séjour déterministe.....	106
5.3.8.7	Temps de séjour stochastique.....	106
5.3.8.8	Coût total	106
5.3.8.9	Pénalité déterministe	106
5.3.8.10	Pénalité stochastique	107
5.3.8.11	Rang Pareto déterministe	107

5.3.8.12	Rang Pareto stochastique	107
5.3.8.13	Évaluation déterministe.....	107
5.3.8.14	Évaluation stochastique.....	107
5.4	Personnalisation du modèle.....	108
5.4.1	Diamètres commerciaux.....	108
5.4.2	Variables.....	109
5.4.3	Contraintes	110
5.4.4	Fonctions objectives	110
5.4.4.1	Coût	111
5.4.4.2	Autres fonctions objectives	111
5.4.5	Extension des scenarios de demande pour la simulation déterministe.....	111
5.4.6	Paramètres de l'algorithme génétique	112
5.4.7	Paramètres d'échantillonnage stochastique.....	112
5.4.8	Modèle de simulation	112
5.5	Conclusion.....	113
CHAPITRE 6	COMPARAISON DES APPROCHES SUR UN EXEMPLE D'APPLICATION	114
6.1	Introduction	114
6.2	Description du système de distribution	114
6.3	Optimisation du système de distribution	117
6.3.1	Approche déterministe	117
6.3.2	Approche stochastique sans corrélations.....	124
6.3.3	Approche stochastique avec corrélations	127
6.4	Conclusion.....	131
CHAPITRE 7	DISCUSSION GÉNÉRALE	133

CHAPITRE 8	CONCLUSIONS ET RECOMMANDATIONS	137
8.1	Conclusions générales	137
8.2	Contributions	138
8.2.1	Articles publiés dans des revues avec comité de lecture.....	138
8.2.2	Articles soumis et en cours de révision	138
8.2.3	Principales contributions à la recherche et au développement.....	139
8.2.3.1	Orientation du choix d'une approche appropriée	139
8.2.3.2	Proposition de paramètres permettant de généraliser la conception	139
8.2.3.3	Développement d'un modèle d'optimisation déterministe et stochastique.....	139
8.3	Recommandations	140
RÉFÉRENCES	142
ANNEXES	150

LISTE DES TABLEAUX

Tableau 1.1 : Justification des méthodes et paramètres utilisés dans la thèse	7
Table 2.1: Some Tank Design Criteria for Low-Density Residential Areas (data from MDDEP 2006; van Zyl et al. 2008a).....	15
Table 2.2: Typical South African Guidelines for Sizing Storage Tanks (Reproduced from Nel, D., and Haarhoff, J. (1996). “Sizing municipal water storage tanks with Monte Carlo simulation.” <i>Journal of Water Supply: Research and Technology-AQUA</i> , 45(4), 203-212, with permission from the copyright holders, IWA Publishing)	17
Table 2.3: Optimization methods application to Rank Design in WDN.....	24
Table 2.4: Strengths and weaknesses of different approaches	26
Table 3.1 : D-Town Pressure zones characteristics.....	53
Table 3.2 : Reservoir volume and pumping capacity required for D-Town network	53
Table 3.3 : Pumping head required for D-Town	55
Table 3.4 : Summary of design values for the seven selected solutions for D-Town.....	58
Table 3.5 : Objective function values found in the Battle compared to solution 4 (data from Marchi <i>et al.</i> 2013).....	59
Table 4.1 : Saint-Jerome Pressure zones characteristics	84
Table 4.2 : Summary of design values for Saint-Jerome network optimization problem.....	88
Tableau 5.1 : Vecteurs de classes élémentaires pour le projet	97
Tableau 5.2 : Système d'unités en fonction du débit	98
Tableau 5.3: Coefficient de conversion du débit en m ³ /s.....	99
Tableau 6.1 : Diamètres commerciaux disponibles	115
Tableau 6.2: Valeurs des fonctions objectives pour les six solutions sélectionnées.....	121
Tableau 6.3 : Résumé des résultats pour le réseau de l'exemple	125

Tableau 6.4 : Calcul de la réserve d'équilibre et du débit de pompage pour le réseau de l'exemple	
.....	128
Tableau 6.5 : Résumé des résultats stochastiques avec corrélation pour le réseau de l'exemple	130
Tableau 7.1 : Résultats obtenus dans la thèse en relation avec les approches et paramètres utilisés	
.....	136

LISTE DES FIGURES

Figure 2.1: Curves for calculating the balancing reserve, (data from Brière 2007).....	14
Figure 2.2: Procedure for selecting appropriate approach	30
Figure 3.1 : Steps overview of the proposed approach	37
Figure 3.2 : Optimization process in the proposed approach.....	48
Figure 3.3 : D-Town network layout.....	50
Figure 3.4 : Storage variation for different districts in D-Town network	54
Figure 3.5 : Genome definition for the optimization of D-Town network	57
Figure 4.1 : Probability density and cumulative probability of Beta symmetric distribution	72
Figure 4.2 : Optimization procedure in the proposed methodology	80
Figure 4.3 : Saint-Jerome WDS layout and elevations	82
Figure 4.4 : Saint-Jerome network diurnal demand pattern	83
Figure 4.5 : Genome definition for the optimization of Saint-Jerome network.....	87
Figure 4.6 : New Tank location and fire flow allocation in the south-east of Saint-Jerome network.....	89
Figure 5.1 : Objets physiques d'un système de distribution d'eau potable (Rossman, 2000)	91
Figure 5.2 : Principe de base de l'Algorithme génétique.....	102
Figure 6.1 : Réseau de l'Exemple (Lansey & Mays, 1989)	114
Figure 6.2 : Frontière de Pareto pour le coût total versus le temps de séjour de l'eau dans l'Alternative 1	119
Figure 6.3 : Frontière de Pareto pour le coût total versus l'indice de résilience dans l'Alternative 1	120
Figure 6.4 : Frontière de Pareto pour l'indice de résilience versus le temps de séjour de l'eau dans l'Alternative 1	120

Figure 6.5 : Frontière de Pareto pour le coût total versus le temps de séjour de l'eau dans l'Alternative 2	122
Figure 6.6 : Frontière de Pareto pour le coût total versus l'indice de résilience dans l'Alternative 2.....	122
Figure 6.7 : Frontière de Pareto pour l'indice de résilience versus le temps de séjour de l'eau dans l'Alternative 2	123
Figure 6.8 : Comparaison des résultats de conception: Lansey & Mays; Solution 2 et Solution 6	124
Figure 6.9 : Comparaison des résultats pour les solutions (0), (2) et (4)	127

LISTE DES SIGLES ET ABREVIATIONS

AADD	Annual Average Daily Demand
AFD	Acre Feet per Day
AHP	Analytic Hierarchy Process
ASCE	American Society of Civil Engineers
AWWA	American Water Works Association
CDF	Cumulative probability Density Function
CFS	Cubic Feet per Second
CMD	Cubic Meters per Day
CMH	Cubic Meters per Hour
DI	Disturbance Index
DM	Decision Maker
DMA	District Metered Area
DOS	Disk Operating System
EA	Evolutionary Algorithm
EPA	Environmental Protection Agency
GA	Genetic Algorithm
GHG	Green House Gas
GPM	Gallons Per Minute
IMGD	Imperial Million Gallons per Day
IPCC	Intergovernmental Panel on Climate Change
ISO	International Organization for Standardization
IWSA	International Water Supply Association
LH	Latin Hypercube

LHS	Latin Hypercube Sampling
LPM	Liters Per Minute
LPM	Liters Per Second
MDDEP	Ministère du Développement Durable, Environnement, Faune et Parcs
MLD	Million Liters per Day
MGD	Million Gallons per Day
MO	Multi-Objective
NSGA	Non-dominated Sorting Genetic Algorithm
NZ	New Zone
PDF	Probability Density Function
PRV	Pressure Reducing Valve
PZ	Pressure Zone
RDEP	Réseaux de Distribution d'Eau Potable
SMGA	Structured Messy Genetic Algorithm
SO	Single-Objective
SPEA	Strength Pareto Evolutionary Approach
US	United States
WDN	Water Distribution Networks
WDS	Water Distribution Systems
WDSA	Water Distribution System Analysis
WTP	Water Treatment Plant

LISTE DES ANNEXES

ANNEXE 1 : Format du fichier d'entrée EPANET.....	150
ANNEXE 2 : Description complète du problème d'optimisation du réseau de D-Town.....	153

CHAPITRE 1 INTRODUCTION

1.1 Contexte de l'étude

Ce projet de recherche se situe dans le contexte global de planification et gestion de la ressource en eau. Cette ressource rare et précieuse est essentielle pour la survie de toute vie sur terre. Le stockage efficace de l'eau et les systèmes de distribution de cette eau sont parmi les plus grands exploits de l'ingénierie, et ils présentent certains des plus grands défis dans le contexte actuel de surpopulation et du réchauffement climatique.

L'optimisation des Réseaux de distribution d'Eau Potable (RDEP) continue d'être un sujet d'actualité vu les défis que présentent la réduction de coût et la construction de réseaux plus fiables par considération d'un large éventail d'objectifs. L'objectif principal de l'optimisation des composantes des RDEP est de minimiser les coûts d'installation (ou de réhabilitation) et d'exploitation de ces infrastructures tout en respectant les exigences sur la pression et la vitesse de l'écoulement dans tout le système. Cependant, au cours des dernières années, on note un intérêt croissant quant à l'obtention d'un compromis entre le coût du système et les bénéfices souvent exprimés en termes de fiabilité du réseau. Ceci a conduit à l'utilisation des techniques d'optimisation multiobjective permettant d'obtenir un ensemble de solutions Pareto-optimal dans l'espace coût-bénéfice. Dans un contexte marqué par le vieillissement des infrastructures de distribution, l'accroissement rapide de la population, l'apparition des normes plus rigoureuses en matière de sécurité, de santé et de protection de l'environnement, plusieurs mesures (quantifications) des bénéfices ont vu le jour. Ces mesures sont de plus en plus intégrées dans la conception optimale des RDEP. Dans cette thèse, certaines de ces mesures de bénéfices seront présentées.

1.2 Optimisation de la conception des réseaux de distribution d'eau potable

Un RDEP est typiquement composé de conduites, réservoirs, pompes et vannes. Les conduites se croisent aux nœuds auxquels sont généralement affectées des demandes en eau considérées comme des débits sortant du réseau. À chaque composante du réseau est associée une élévation. Les réservoirs sont caractérisés par un volume d'eau en stock, une énergie potentielle et un débit.

Initialement, le réseau était conçu de telle sorte que le problème d'optimisation consistait simplement au choix des caractéristiques des composantes pour une configuration statique de l'ensemble des composantes du réseau. Par la suite, il y'a eu plusieurs tentatives d'inclusion de l'emplacement de certaines composantes dans le processus d'optimisation, ce qui complique considérablement le problème (Afshar et al. 2005).

L'optimisation procède en considérant des tailles alternatives et des politiques d'opération pour les éléments du système et pour chaque configuration du réseau en calculant les propriétés hydrauliques du réseau telles que les valeurs de pression et vitesse. Le calcul des propriétés hydrauliques du réseau est communément connu par le terme « équilibrage du réseau » et constitue à lui seul un problème complexe. Le système a une configuration faisable si les propriétés hydrauliques satisfont aux contraintes qui leur sont imposées. Cependant, il est également possible d'envisager des configurations non faisables en fonction du degré de violation des contraintes. Dans une recherche exhaustive de configurations, chaque élément du système prendrait chacune de ses valeurs possibles, générant ainsi de multiples combinaisons. Ces combinaisons augmentent de façon exponentielle lorsque le nombre d'éléments du réseau augmente; ce qui donne lieu à une explosion de l'espace de recherche.

Le diamètre d'une nouvelle conduite est typiquement un des diamètres commerciaux disponibles, chacun étant associé à un coût différent. En plus, les conduites existantes dans un réseau en réhabilitation peuvent subir plusieurs opérations notamment le nettoyage, la duplication, le remplacement ou l'élimination. Les réservoirs sont utilisés comme éléments d'équilibrage pour combler la demande pendant la pointe et emmagasiner le volume d'eau excédentaire pendant la période de faible demande. Ces réservoirs aident également à la stabilisation des pressions dans le réseau. Ils peuvent être placés à différents points dans un système. Leurs coûts d'installation varient selon leurs tailles. Les pompes peuvent également être placées dans un système pour ajouter de l'énergie au besoin. Il existe différents types de pompes, chacune avec des coûts d'installation et d'exploitation différents. Les pompes sont également associées à des courbes caractéristiques et de rendement selon le débit pompé et l'énergie fournie. Les vannes sont utilisées pour modifier les profils d'écoulement dans les conduites et le plus souvent pour réduire la pression entre les différentes parties du système. L'optimisation s'applique typiquement à un système en état permanent dans les conditions de pointe, bien que plusieurs conditions de chargement puissent être analysées. Il est supposé que

les entrées et les sorties du réseau sont connues à l'avance. Si les réservoirs sont inclus dans la conception, il devient alors essentiel de procéder à une simulation de longue durée des entrées et sorties des réservoirs pour une conception qui donne lieu à une politique d'opération efficace des réservoirs (Walski et al., 2000).

Étant donné une configuration de réseau, il existe des modèles de simulation hydraulique capables de calculer les propriétés hydrauliques de son état. Le modèle de simulation très connu EPANET (Rossman, 1999) sera utilisé dans cette thèse à cette fin.

1.3 Motivation de la recherche

L'exigence de l'humanité et l'évolution de la technologie a permis d'améliorer les réseaux de distribution au fil des années. Compte tenu de l'importance capitale que représente l'eau pour l'humanité, sa distribution doit faire partie des objectifs principaux de toute administration. L'extension et la réhabilitation des réseaux existants constituent déjà une priorité de plusieurs services de distribution d'eau dans le but de répondre à la demande future et d'améliorer le service rendu à la population. L'amélioration du service inclut le contrôle de la qualité de l'eau distribuée pour réduire les risques d'épidémie. Toutes ces considérations complexifient les RDEP qui comprennent aujourd'hui des stations de traitement d'eau, des réseaux complexes, des stations de pompage et des réservoirs d'eau. De nombreux défis continuent d'être posés aux gestionnaires des réseaux, vue l'augmentation sans cesse des coûts unitaires de distribution d'eau aux populations. Ces coûts sont liés à la conception, l'exploitation et la maintenance. A titre d'exemple, les RDEP consomment dans certains pays industrialisés jusqu'à 20% de l'électricité utilisée par les collectivités publiques (InfraGuide, 2005). La maintenance nécessite parfois l'ajout de nouveaux réservoirs qui jouent plusieurs rôles (chapitre 2). La conception optimale de ces réservoirs permet de faire des économies, d'augmenter la fiabilité du réseau et d'améliorer la qualité de l'eau distribuée.

Dans plusieurs pays et régions, l'approche utilisée pour la conception des réservoirs dans les RDEP est basée sur des guides techniques. Au Québec par exemple, ces règles sont dictées par les compagnies d'assurance et le ministère de l'Environnement (MDDEP, 2002). L'inconvénient dans ces approches est de faire appel à de grands réservoirs qui peuvent être hydrauliquement nécessaires mais défavorables quant à la qualité de l'eau suite à de longs temps de résidence. Les

réservoirs ont déjà été la cible de plusieurs recherches à l'École Polytechnique de Montréal dans les années antérieures notamment à la chaire en eau potable (Basile, 2008; Chevalier, 2002; Gauthier et al., 2000). L'optimisation des réservoirs permettrait d'améliorer les problèmes de qualité relevés dans ces études. La disponibilité d'outils informatiques performants rend cette tâche possible et permet de prendre en considération plusieurs paramètres tels que la qualité de l'eau, la performance hydraulique, le coût de construction et d'exploitation lors de la conception.

1.4 Hypothèses scientifiques de la recherche

Les hypothèses sur lesquelles s'appuie cette recherche sont :

- Hypothèse 1 :

Il existe un certain écart susceptible d'être réduit entre les approches proposées dans la littérature et celles utilisées par les ingénieurs sur le terrain pour la conception des réservoirs dans les réseaux de distribution d'eau potable. L'objectif spécifique 1 vérifie cette hypothèse.

- Hypothèse 2 :

L'utilisation d'une approche généralisée de formulation du problème d'optimisation multiobjective des réseaux, couplée avec un algorithme génétique pour la résolution, permet une conception efficace qui puisse s'adapter aux exigences du terrain et s'appliquer sur un réseau complexe et réaliste, permettant ainsi de réduire l'écart avec les approches de terrain. L'objectif spécifique 2 permettra de vérifier cette hypothèse.

- Hypothèse 3 :

L'implémentation de l'optimisation de la conception des réseaux dans un environnement stochastique intégrant les aspects de la qualité de l'eau est plus représentative de la réalité et plus efficace pour la conception des réseaux réels de distribution d'eau potable. L'objectif spécifique 2 permettra de vérifier cette hypothèse.

- Hypothèse 4 :

La prise en compte des corrélations entre les demandes aux nœuds dans l'optimisation stochastique des réseaux de distribution conduit à une conception plus coûteuse mais aussi plus fiable. L'objectif spécifique 3 permettra de vérifier cette hypothèse.

Les aspects originaux de la recherche sont : (1) la classification structurée des approches avec leurs avantages et inconvénients et l'orientation de choix d'une approche adaptée; (2) la proposition d'étapes guidées à suivre lors de l'optimisation des réseaux de distribution; (3) le développement d'un outil d'aide à la décision dans la conception et la gestion des réseaux de distribution ainsi que (4) la prise en compte des aspects de la qualité de l'eau dans l'optimisation en environnement stochastique des réseaux de distribution d'eau potable.

1.5 Objectifs de la recherche

Afin de garantir une durabilité prolongée des infrastructures et une distribution efficace et peu coûteuse de l'eau potable, la conception de nouveaux réseaux de distribution ou leur réhabilitation nécessite la prise en compte d'aspects qui vont au delà du coût. Le présent travail de recherche doctorale s'inscrit dans le cadre d'un projet dont l'objectif à terme est le développement d'un modèle d'optimisation multiobjective des réseaux de distribution d'eau potable capable de prédire la nécessité d'introduire un ou plusieurs réservoirs dans le réseau ainsi que les emplacements de ces éventuels réservoirs. Le souci de jalonner le chemin conduisant à l'objectif principal ci-dessus énoncé a conduit à l'élaboration d'une stratégie qui se décline en trois étapes ou objectifs spécifiques :

1. Analyser et comparer minutieusement les différentes approches utilisées pour la conception et la gestion des réservoirs dans les réseaux de distribution d'eau potable à travers une revue critique de la littérature.
2. Développer un modèle d'optimisation multiobjective à deux variantes (l'une déterministe et l'autre stochastique) basée sur un algorithme génétique pour la conception et la gestion des réseaux de distribution d'eau potable incluant les réservoirs et en évaluer les performances.
3. Faire une étude comparative des différentes approches sur un exemple de réseau dans le but de fournir aux gestionnaires des municipalités un outil d'aide à la décision.

1.6 Plan de la thèse

Outre le présent chapitre introductif, la rédaction de la présente thèse s'articule autour de cinq principaux chapitres :

- Le chapitre 2 propose une revue critique de la littérature sur la conception optimale des réservoirs dans les réseaux de distribution d'eau potable.
- Le chapitre 3 présente le modèle d'optimisation multiobjective déterministe permettant la conception ou la réhabilitation d'un réseau de distribution d'eau potable.
- Le chapitre 4 présente une implémentation novatrice du problème de conception des réseaux en environnement stochastique en incluant la qualité de l'eau dans les objectifs afin d'étendre efficacement la conception aux réservoirs.
- Le chapitre 5 présente certains détails du modèle développé tout en fournissant une orientation dans la personnalisation des variantes du modèle.
- Le chapitre 6 étudie l'impact de différentes approches de conception sur le coût et la fiabilité en appliquant ces approches sur une étude de cas.

Cette étude se termine par une discussion générale ainsi qu'une conclusion et des recommandations visant à souligner les limites du modèle développé et proposer les possibilités d'améliorations futures.

1.7 Justification des méthodes et paramètres utilisés dans la thèse

Plusieurs décisions ont été prises quant aux méthodes et paramètres à utiliser à diverses étapes de la mise en œuvre de la thèse. Le Tableau 1.1 donne une synthèse des justifications de ces décisions.

Tableau 1.1 : Justification des méthodes et paramètres utilisés dans la thèse

Étape	Description		Décision	Justification
Formulation	Type		Multiobjective	Proposition de solutions alternatives donne plus de choix aux concepteurs pour faire un compromis selon leurs priorités; réaliste et adapté au problème de conception des réseaux
	Indicateur de la qualité de l'eau		Temps de séjour	Le temps de séjour est utilisé pour calculer d'autres indicateurs de qualité tels que la croissance de microorganismes; plus le temps de séjour est court, meilleure est la qualité de l'eau
Optimisation	Méthode		Algorithme Génétique	Adaptée pour des problèmes faisant intervenir des variables discrètes; moins sensibles aux paramètres utilisés; fiable et largement adoptée dans la littérature pour la conception des réseaux de distribution
Simulation	Modèle		EPANET	Robuste et libre d'accès; possibilité d'échanger les données avec Autocad et ArcGis facilite la préparation du réseau; simule la qualité de l'eau en plus de la simulation hydraulique
	Approche		Déterministe et Stochastique	Comparer deux approches de conception des réseaux en identifiant les avantages et les inconvénients de chacune afin d'orienter le choix d'une approche appropriée
	Demande aux noeuds	Méthode d'échantillonnage	Latin Hypercube	Représentativité de chaque intervalle le long de la palme de variation des demandes; permet d'utiliser un nombre faible d'échantillons pour une simulation stochastique fiable
		Loi de distribution	Beta	Bornée donc permet à l'utilisateur de choisir sa plage de variation des demandes; évite la génération de demandes extrêmes irréalistes
		Dépendance	Absence et existence de corrélations	Évaluer l'impact des corrélations sur les solutions proposées afin de déterminer le niveau de sensibilité du modèle vis-à-vis de l'hypothèse sur la dépendance des demandes aux noeuds
Programmation	Langage		C++	Rapidité d'exécution; flexibilité et bonne gestion de l'allocation de la mémoire; compatibilité avec la librairie de l'algorithme génétique et le code source d'EPANET
Validation	Nombre d'études de cas		3	Pour tester différents aspects de l'approche proposée pour l'optimisation de diverses composantes du réseau sur plusieurs types de projets de conception: nouveau réseau, réhabilitation et extension

CHAPITRE 2 ARTICLE 1: OPTIMAL TANK DESIGN IN WATER DISTRIBUTION NETWORKS: REVIEW OF LITERATURE AND PERSPECTIVES

2.1 Présentation de l'article

Dans cette section, l'objectif de l'article sera présenté, ainsi qu'une conclusion sur ce travail dans le cadre de la thèse. Cet article présente une revue comparative et critique des approches de dimensionnement des réservoirs, comme partie intégrante des réseaux de distribution d'eau potable (RDEP). Il peut être consulté par les acteurs de la gestion des RDEP à la recherche des idées préliminaires sur l'approche à utiliser pour intégrer les réservoirs dans la conception de leurs réseaux.

2.1.1 Objectif

Lorsqu'un RDEP contient un ou plusieurs réservoirs, la conception et l'opération de l'ensemble du réseau dépend essentiellement des décisions prises sur les réservoirs telles que leur nombre, leur taille ou leur emplacement. Une mauvaise conception de ces réservoirs dans les réseaux entraîne alors des conséquences globales néfastes sur les performances et voire même sur le coût de l'ensemble du réseau. Étant donné que dans cette thèse, on s'intéresse à une conception optimale des réservoirs, il s'avère indispensable de faire une analyse critique de ce qui est proposé dans la littérature ou ce qui se pratique actuellement dans la conception des réservoirs avant de proposer une approche de conception qui étend l'analyse aux autres composantes du réseau. Pour atteindre cet objectif, l'approche actuelle utilisée par les ingénieurs sur le terrain ainsi que les modèles d'optimisation proposés dans la littérature sont analysés. Les origines de l'écart entre les deux approches sont examinées dans le but de s'orienter vers une approche qui tend à réduire cet écart. Une procédure est enfin suggérée pour guider au choix d'une approche appropriée sur la base des données disponibles.

2.1.2 Conclusion

On constate d'une part que la plupart des modèles n'intègrent pas les réservoirs dans le processus d'optimisation des RDEP. D'autre part, les modèles qui intègrent les réservoirs posent

généralement plusieurs hypothèses simplificatrices pouvant affecter la qualité et l'utilité des résultats. Par ailleurs, l'approche conservative basée sur l'expérience en génie et les guides techniques s'est révélée la plus répandue en matière de conception des réservoirs dans les RDEP. L'écart important observé entre cette approche et les modèles mathématiques d'optimisation explique en partie la non-adoption de ces modèles par les ingénieurs praticiens. Parmi les origines de cet écart, on a noté le manque de collaboration entre les chercheurs et les ingénieurs du terrain pour partager les vrais enjeux des décisions à prendre. La nature stochastique du problème lui-même explique aussi en partie les divergences observées. L'utilisation des algorithmes génétiques dans une approche multi-objective avec prise en compte des incertitudes liées aux données et la validation sur des études de cas réels peut réduire l'écart et améliorer l'applicabilité des modèles d'optimisation par les ingénieurs praticiens.

Cet article est publié à "Journal of water resources planning and management".

Batchabani, E. and Fuamba, M. (2014b). "Optimal Tank Design in Water Distribution Networks: Review of Literature and Perspectives." *Journal of Water Resources Planning and Management*, 140(2), 136–145.

2.2 Optimal tank design in water distribution networks: review of literature and perspectives

Essoyeke Batchabani¹ and Musandji Fuamba, Ph.D.²

¹Ph.D. Candidate, Dept. of Civil, Geological and Mining Engineering, École Polytechnique de Montréal, 2500 Chemin de Polytechnique, P.O. Box 6079, Station Centre-Ville Montréal, Montréal, QC, Canada H3C 3A7 (corresponding author). Email: essoyeke.batchabani@polymtl.ca

²Associate Professor, Dept. of Civil, Geological and Mining Engineering, École Polytechnique de Montréal, 2500 Chemin de Polytechnique, P.O. Box 6079, Station Centre-Ville Montréal, Montréal, QC, Canada H3C 3A7. Email: musandji.fuamba@polymtl.ca

2.2.1 Abstract

Water Distribution Networks (WDN) are traditionally designed without integrating tanks in the optimization process, despite the numerous benefits tanks bring to network operation. Since 2004, a number of efforts have been made to integrate tanks, and especially consideration of tank sizing, in optimization models for WDN, but important limitations in these models have impeded their adoption. This paper reviews different approaches to including tank consideration in WDN optimization, and examines the gap between proposed models and current engineering practice. It then analyzes the strengths and weaknesses of these models and suggests a procedure for selecting the most suitable approach to real-world WDN optimization projects. Further research is proposed to improve the validity and applicability of models to decision-makers and policymakers.

Keywords: Water distribution networks, Water tank or reservoir, Optimization, Multiobjective.

2.2.2 Introduction

Storage tanks play an important role in water distribution networks (WDN) (Brière, 2007), enabling demand management, assuring water supply in case of system failure and reserves for emergencies such as fire fighting, and allowing for the modulation of pump flow rate. When

properly designed and located, storage tanks are a cost-effective means of improving overall network performance. However, an inadequately designed or located tank can increase the cost of pipe design and pump operation, and reduce network performance indices such as reliability and resilience (Vamvakeridou-Lyroudia et al. 2007).

WDN optimization models that have been used, including linear programming (Alperovits & Shamir, 1977; Fujiwara, 1990; Quindry et al. 1979; Quindry et al. 1981; Sherali et al. 1998), techniques for total/partial or implicit/explicit enumeration (Savic & Walters, 1997), and nonlinear programming techniques (Djebedjian et al. 2000), focus solely on pipe diameter optimization, without consideration of tank location and design (Farmani et al. 2004), which are decided on the basis of land availability and expert opinion.

Tank design denotes here the set of decisions regarding the characteristics of tanks used in the network. This paper focuses on the hydraulic aspects of design as opposed to structural, corrosion, safety, contamination, or instrumentation aspects. Selecting a tank design for a given WDN requires consideration of design characteristics such as (1) supply volume (balancing, fire, emergency), (2) hydraulic (bottom elevation, top level, mixing behavior, retention time) (3) operational (bottom storage, free volume, maximum operational level and minimum normal operational level) and (4) construction (type - buried/elevated, shape, location, position and configuration of the inlet and outlet pipes). These design characteristics must be balanced to maximize reliability and water quality, and minimize capital and operational costs. All these variables need to be considered as decision variables in order to incorporate tank design into a global WDN optimization model. This is complex: while it is not recommended to consider decision variables with either quasi-constant values, or with little or no effect on the objective function value (little benefit), if too many variables are ignored the optimization algorithm can produce solutions which would be mathematically possible but technically and commercially inapplicable (Vamvakeridou-Lyroudia et al. 2005). Given these limitations, current models make assumptions and simplifications to incorporate tanks in WDN optimization. Some assign a fixed shape or a fixed ratio for emergency volume in relation to total volume, while others consider only elevated tanks or omit variables such as risers. These assumptions affect the quality and usefulness of results.

Given the growing recognition of the value tanks add to WDN, analysis of different means for incorporating tank considerations into WDN optimization models is needed. This paper seeks to contribute to this effort and thereby improve decision-making and planning in this area. The paper (1) reviews current approaches to tank design, including optimization models and the selection of tank design variables; (2) examines the gap between engineering practice and optimization models in designing tanks for WDN; (3) provides an analysis of the strengths and weaknesses of current approaches; (4) details a step-wise procedure for selecting the best approach for a given WDN project and (5) identifies prospects for future research.

2.2.3 Current practice to tank design in WDN

2.2.3.1 Type

In current practice, the choice of tank type is influenced by factors such as topography, potential for freezing, aesthetics, land availability, seismic considerations, cost and experience. Elevated tanks are generally preferred for equalizing storage in a water distribution system. In fact, once installed, elevated equalizing storage becomes a control on the hydraulic performance of a system (McPherson, 1966). Buried tanks may be preferred in cold climates because they are less susceptible to freezing. However, elevated tanks could be used in these climates, provided they are heated (Hodnett, 1981). Elevated tanks and standpipes are needed in flat areas. In hilly areas, it is possible to place ground storage on hills at low cost (Walski, 2000).

2.2.3.2 Location

Tank location is generally determined by land availability. Topography and aesthetics are also considered. Hills are good locations for tanks since ground tanks can act as elevated ones in such locations. In flat areas, there is usually more flexibility with regard to location. It is common practice to locate tanks near high-demand areas (Mays, 2000), as this reduces storage capacity requirements (Piqueiro et al. 1991). According to Walski (2000), storage tanks are best placed on the downstream side of the area of largest demand, as this allows the use of smaller pipe sizes and reduces interruptions caused by pipe failure.

2.2.3.3 Levels

The key point in determining tank elevation is to identify the elevation of the pressure zone the tank will serve. The highest and lowest users in the zone determine if it is possible to get by with one zone or if several pressure zones are required. This information, combined with the minimum and maximum allowable pressures, determines the tank levels (overflow, minimum normal day and bottom of tank). The highest customer receives the minimum allowable pressure when the tank is at its minimum normal level. The lowest customer receives the maximum allowable pressure when the tank is full and there is little head-loss in the system (Walski, 2000).

2.2.3.4 Volume

Tank volume is actually determined by standards or guidelines specific to each governmental jurisdiction. These guidelines usually focus on tank volume and not on factors such as bottom and top elevation or location. They are based on the amount of storage water required for various usages, such as the balancing storage (R_{bal}), fire storage (R_{fire}), emergency reserve (R_{emerg}) and production reserve (R_{prod}). According to Brière (2007), the balancing storage for a given day represents about 20% of daily consumption, the maximum balancing volume being calculated based on daily maximum consumption. The plot of the cumulative daily consumption curve used to estimate the balancing reserve is shown in Figure 2.1, which is obtained by (1) plotting the peak day's consumption curve (demand flow); (2) plotting the peak day's cumulative consumption curve (cum. demand); (3) plotting a straight line from the point representing cumulative consumption at one hour to the point representing cumulative consumption at 24 hours (uniform pumping flow). The slope of this line represents the average consumption rate for that day and this line can be interpreted as a uniform pumping flow rate which would supply the same cumulative volume in a 24-hour period. The next step (4) involves plotting two straight lines parallel to the line plotted above: one is tangential to the upper part of the cumulative consumption curve (upper tangent), corresponding to an empty tank, and the other is tangential to the lower part of that curve (lower tangent), corresponding to a full tank. (5) The vertical distance between the two parallel lines represents the balancing reserve (Br). These methods have generally been replaced by extended period simulation models which can be more flexible than the graphical methods.

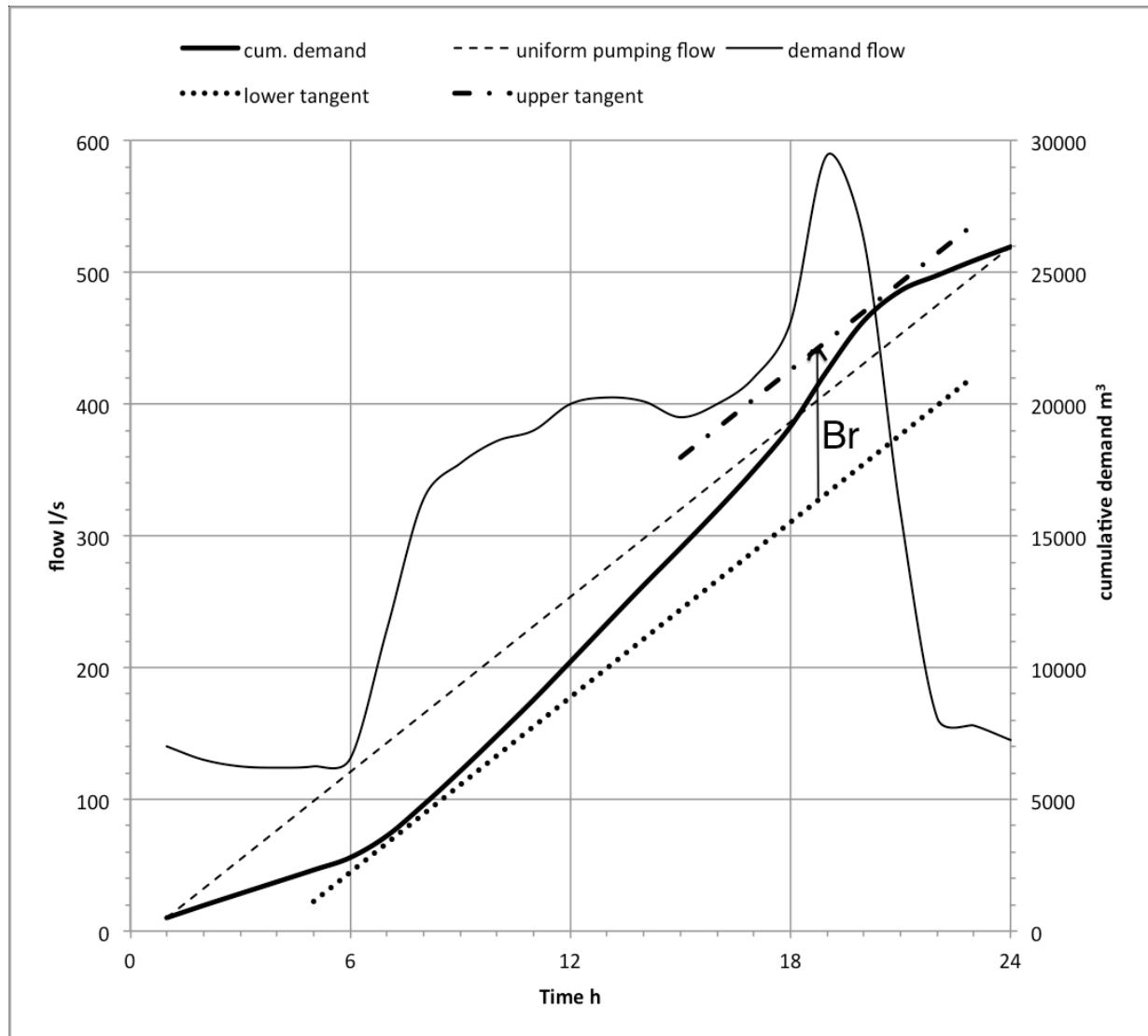


Figure 2.1: Curves for calculating the balancing reserve, (data from Brière 2007)

In some countries, water storage requirements for fire protection may be generated by fire-insurance underwriters. The formulae to determine the required fire flows depends on the construction type and the total floor area (AWWA, 2008; Fire-Underwriters-Survey, 1999). The emergency reserve must be sufficient to meet average daily consumption for 2 to 14 hours, and the production reserve must sustain approximately 4 production hours. This production reserve is usually coupled with the treatment plant and stored in a separate reservoir. Since the probability of having simultaneously (a) a maximum consumption (R_{bal}), (b) a major fire (R_{fire}), (c) failure of one of the main network pipes (R_{emerg}), and (d) a breakdown at the treatment plant (R_{prod}), is very low, it is not economically justifiable to store a volume of water equal to the sum of all the water

reserves at all times (Brière, 2007). The maximum (R_{max}), minimum (R_{min}) and desired (R_{desir}) reserves are then determined by equations (2.1) through (2.3).

$$R_{max} = R_{bal} + R_{fire} + R_{emerg} + R_{prod} \quad (2.1)$$

$$R_{desir} = R_{bal} + R_{fire} + R_{emerg} \quad (2.2)$$

$$R_{min} = R_{bal} + R_{fire} \quad (2.3)$$

Table 2.1: Some Tank Design Criteria for Low-Density Residential Areas (data from MDDEP 2006; van Zyl et al. 2008a)

Country	Storage requirements		
	Balancing (B)	Fire (F) (m ³)	Emergency (E) (days)
United States	35 % of peak day	340 ^a	1 peak
France	30 % peak day	120	2 average
South Africa	Included in E	108	2 average
Quebec	Included in E	120	1.5 average

^aUpper limit based on the ISO evaluation method

Design guidelines specifying the tank volume for balancing, fire and emergency storage for low-density areas in the United States, France, South Africa and Quebec are illustrated in Table 2.1 (van Zyl et al. 2008 ; MDDEP, 2006). Fire storage of 340 m³ in Table 2.1 for United States is an upper limit based on the ISO evaluation method, which uses: fire storage required = fire demand + peak day demand – production capacity. Fire demand for low-density residential areas is usually 750 gpm (2840 l/min) for 2 hours. Since production should exceed demand, the 340 would be an upper limit. The three columns in Table 2.1 are not systematically additive since the design fire is unlikely to occur at the same time as a major outage. But most guidelines specifically require these values to be added together to obtain the tank volume (as in equation (2.2)). Larger volumes can be justified in areas subject to time of day energy pricing so that tanks can be filled when energy is cheap and pumps turned off during high-energy use hours. The requirements for tank volume in Quebec are less stringent than in the United States and France. European countries often have substantially lower fire water requirements than the

United States, mainly due to water quality concerns (van Zyl & Haarhoff, 1997). Quebec's tank volume requirements are based on MDDEP (Ministère du Développement Durable, de l'Environnement et des Parcs) regulations that estimate volume on the basis of average daily consumption, assuming that the supply source is able to produce the daily peak demand during the planned period. In practice, the current daily consumption average is increased by a factor to take future demand into account. This factor depends on the growth rate in the service area and design should be based on system demand at the end of the study period. Some engineers in Quebec use a factor of 10% as a rule of thumb. Indeed no study has yet been conducted to support the 10% value, which is derived from current practice in Quebec, as reported by a practicing engineer with 30 years experience (Pierre Ethier 2012, personal communication). The sum of the operational (balancing and emergency) and fire storage represents the useful storage of the tank. A given country may have different guidelines for tank volumes depending on the authority responsible for the development in question. Table 2.2 shows tank storage requirements in the guidelines of different jurisdictions in South Africa. In this Table, AADD means annual average daily demand.

Current guidelines give the design engineer and regulators considerable discretion in the design (Walski, 2000). However, surveys presented at the 1955 conference of the IWSA (International Water Supply Association) showed that depending on guidelines, the reservoir capacity can range from less than 50% to more than 200% of the maximum daily production at the treatment plant (IWSA, 1955). The continued use of these storage guidelines despite the large variation in capacity guidelines reveals the inherent conservatism in this area. Although tank size ought to be site specific, design guidelines are conceived for a wide range of systems. Some storage tanks are therefore likely much larger than they need to be.

The limitations of the guidelines approach to tank design highlight the need for an alternative approach to design improvement to support decision-making in tank design optimization in WDN, especially in the actual context of increasing water demand where consumers have new water needs to ensure their comfort. This alternative approach should include application of the optimization methods and integrate site-specific constraints to handle the problem of tank sizing. Since optimization remains a big challenge for engineers in their daily decisions, analyzing

alternatives enables them to make rational decisions that, while not necessarily optimal, are acceptable within the set of solutions.

Table 2.2: Typical South African Guidelines for Sizing Storage Tanks (Reproduced from Nel, D., and Haarhoff, J. (1996). “Sizing municipal water storage tanks with Monte Carlo simulation.” *Journal of Water Supply: Research and Technology-AQUA*, 45(4), 203-212, with permission from the copyright holders, IWA Publishing)

Authority	Nature of supply	Storage tank capacity
Department of Water Affairs	gravity feed	24 h of AADD
	pumped main	48 h of AADD
Co-operation & Development	gravity feed	24 h of AADD
	pumped main	48 h of AADD
National Building Institute	one source	48 h of AADD
	two sources	36 h of AADD
National Housing	gravity feed	20 h of AADD
	pumped main	30 h of AADD
	two sources	66% of capacity with one source

2.2.4 Optimization methods application

Optimization methods coupled with simulation models could improve the design of large-scale networks. In some cases, optimization methods even offer potential cost savings. Wu *et al.* (2005) conducted a study to identify system improvement alternatives in order to satisfy the increasing demand in a growing community. Their study shows that optimization modeling is a powerful approach for supporting and enhancing a sound decision-making process in water industry. Optimization also proved useful for operational control at the pumping station (Klempous et al.1997). Optimization modeling tools are especially helpful when water supply

faces multiple constraints and practical engineers and decision-makers must make a great number of decisions.

Optimization methods for WDN tank design involve a four-part process: selection of the reservoir decision variables, formulation of the problem, resolution of the formulated problem, and validation of the model. Validation is possible through case study.

Formulation of the optimization problem is considered as follows:

$$\min f(V) \text{ and/or } \max g(V) \quad (2.4)$$

subject to:

$$\text{continuity, conservation of energy and design criteria} \quad (2.5)$$

The variables involved in optimization are represented by V in equation (2.4). Design criteria in equation (2.5) include minimum and maximum tank levels. Equation (2.4) gives rise to two types of model. The first is known as the single-objective model and uses cost as the only optimization criterion. The second is known as the multiobjective model and seeks to maximize benefit as well as minimize cost.

Walski (2001) objects to the use of cost minimization as a WDN design objective, firstly because it is difficult to define constraints and benefits due to uncertainty over future demand and secondly because there will always be several different ways of achieving the similar costs and benefits given the range of uncertainty.

2.2.4.1 Single-Objective Models

The general formulation of single-objective models for WDN that include tanks is given in equations (2.6) through (2.9) where n is the number of nodes; m , the number of tanks; r , the number of pumps; q , the number of pumping periods; C the cost; D and L , the pipe diameter and pipe length; V_r , the tank variables (volume, type and shape); V_p , pump variables; P_p , pumping period; H , the head; and A the commercially-available pipe diameters.

$$\min f_1 = \sum_{i=1}^n C_i(D_i, L_i) + \sum_{j=1}^m C_j(V_r) + \sum_{k=1}^r C_k(V_p) + \sum_{l=1}^q C_l(P_p) \quad (2.6)$$

$$g_j(H, D) = 0 \quad (2.7)$$

$$H_j \geq H_{j,min} \quad (2.8)$$

$$D_i \in \{A\} \quad (2.9)$$

Other criteria may also be considered. For instance, in the extended period simulation, tank capacity requires that the inflow volume V_{in} equal the outflow volume V_{out} .

$$\sum V_{in} = \sum V_{out} \quad (2.10)$$

Single-objective optimization has long been used in WDN design. Some models consider only one design parameter for a tank. In the single-objective model proposed by Alperovits and Shamir (1977), for instance, the only decision variable for a tank is the elevation at which it is to be located. The objective function, to be minimized, reflects the overall capital cost plus present value of operating costs. Tank cost was approximated as a linear function of the water level in the reservoir. The model by Ostfeld (2005) defines only tank maximum storage volume. The objective is to minimize the total cost of designing and operating the system for the operational time horizon of 24h. Tank construction cost is a function of maximum volume of water stored in the reservoir at the end of given time steps. The approach, described by Murphy *et al.* (1994) assumes that tanks are conceived to regularize pumping, and uses demand balancing flows in tank design. The location of the tank, its volume and minimum operational level are treated as decision variables. Tanks are considered as nodes with known flow rates. For each hydraulic time step, the balancing flow is calculated in proportion to the trial tank's volume, and is assigned to the corresponding nodes. Tank dimensions are obtained through steady-state hydraulic analyses performed at each time step. Tank costs are a function of the elevation and the volume of storage, and include land acquisition costs. Another similar approach with single objective function involves the use of the penalties function for tank emergency volume checks (Oliveira Sousa et al. 2005). The single-objective model proposed by Prasad (2007) includes five tank decision variables: location, volume, minimum operational level, height-diameter ratio, and emergency-volume-total-volume ratio. This approach assumes that each new tank is cylindrical in shape and that the free volume represents 10% of the total volume. Tank characteristics can be determined after identifying each of these variables. Prasad (2010) proposed another single-objective model for tank design that does not require explicit consideration of operational constraints such as limits on maximum and minimum tank levels, but does require thorough determination of tank design parameters: the location, diameter, elevation, minimum and maximum operating levels, and free-surface height of each tank must be clearly defined. This

can be achieved by encoding such parameters either explicitly or implicitly in a genetic algorithm chain.

Explicit encoding requires extra care to ensure that the tank's bottom elevation is lower than its operational level. Implicit encoding (Prasad, 2010) considers the following decision variables for each new tank: location, total volume, minimum operating level, ratio between diameter and height, and ratio between emergency volume and total volume. Design parameters are calculated using those values and are assigned to tanks connected at trial locations. The main aim of Prasad's model was to demonstrate the importance of reducing operational constraints and performing extended-period simulations as part of the optimization process. Tank designs with the least pressure violations and fewest constraint violations are then retained and ranked according to cost. Minimization of total cost, which includes capital cost and energy cost, is considered as the objective. Tank capital cost is a function of tank size (volume) and the cost of any intermediate tank size is obtained by interpolating linearly from the cost of given standard sizes.

2.2.4.2 Multiobjective Models

Multiobjective optimization models were developed in part because they consider factors such as cost and reliability that figure prominently in the decision making of water utilities.

These models start with the formulation expressed in equations (2.6) to (2.10), and add consideration of benefit (P_r) which is a function of the design variables represented by C_{onc} , as expressed in equation (2.11).

$$\max f_2 = P_r(C_{onc}) \quad (2.11)$$

This benefit is a measure of network reliability which, in WDN, refers to the network's ability to supply consumers adequately under normal and exceptional conditions (Savic et al. 2008). WDN reliability can be assessed through two types of failures, mechanical and hydraulic (Farmani et al. 2005). Mechanical failure refers to component failures such as broken pipes or defective pumps. Hydraulic failure can result from the randomness of certain variables such as consumer demand, pipe roughness, tank levels and supply source levels.

Benefit should be defined in the design of WDN that include tanks. Tank size influences both mechanical and hydraulic failure. For instance, if the tank elevation is too high, the pressure in pipes will also be high, increasing the likelihood of pipe breakage. On the other hand, if the tank elevation is set too low, the water pressure delivered to consumers may be insufficient, resulting in hydraulic failure. Each failure leads to internal water losses and network malfunction. However, these failures can be overcome if the network is designed to have enough surplus energy for internal dissipation (Prasad, 2004). The resilience index I_r was developed as a measure of reliability to account for means taken to mitigate the risk of failures.

The concept of resilience was introduced by Todini (2000) as a way of measuring WDN reliability. This resilience index is defined by equation (2.12) where P_{int} is the network power required to meet total demand and $P_{max,int}$ is the maximum power that can be dissipated internally to meet demands and maintain pressure at the various nodes.

$$I_r = 1 - \left(\frac{P_{int}}{P_{max,int}} \right) \quad (2.12)$$

$$P_{int} = P_{tot} - P_{ext} \quad et \quad P_{max,int} = P_{tot} - P_{min,ext} \quad (2.13)$$

With

$$P_{tot} = \gamma \sum_j Q_j H_j + \sum_k P_k \quad (2.14)$$

$$P_{ext} = \gamma \sum_i q_i h_{disp,i} \quad (2.15)$$

$$P_{min,ext} = \gamma \sum_i q_i h_{nec,i} \quad (2.16)$$

Where P_{tot} is the total available power; P_{ext} the power delivered to the user in terms of flow (q_i) and head ($h_{disp,i}$) at each node i ; $P_{min,ext}$ is the minimum power to be delivered at the minimum required head $h_{nec,i}$; γ the specific weight of water; Q_j and H_j are the flows and the heads of tanks j ; P_k , is the power introduced by pump k .

Given equations (2.14) to (2.16), I_r can be rewritten as in equation (2.17).

$$I_r = \frac{\sum_i q_i (h_{disp,i} - h_{nec,i})}{\left(\sum_j Q_j H_j + \sum_k \frac{P_k}{\gamma} \right) - \sum_i q_i h_{nec,i}} \quad (2.17)$$

In this expression of benefit, equation (2.11) becomes:

$$\max f_2 = I_r \quad (2.18)$$

The multiobjective model developed by Walters *et al.* (1999) specifies tank variables that need to be considered in assessing reliability. These include location, volume and water levels. Location and volume can be considered as independent variables by genetic algorithm (Murphy *et al.* 1994). The two objectives in Walters' model are minimum cost and maximum benefit. Tank costs are considered as a function of volume. The benefit is treated as the reduction in predicted hydraulic deficiencies.

The model by Farmani *et al.* (2004) uses advanced tank sizing. In fact in this model, the control variables are pipe rehabilitation decisions, tank design, tank location, and pumping schedule. Farmani's system is designed to operate under multiple loading conditions. A maximum of two new cylindrical tanks can be considered in this particular modeling approach. Tank location, normal minimum/maximum daily level, tank diameter and bottom level are considered independent variables. The first objective of the model is to minimize cost. Tank costs are again considered as a function of volume. The resilience index is considered as a second objective to increase the hydraulic reliability and the availability of water during network failures.

The model proposed by Vamvakeridou-Lyroudia *et al.* (2005) considers tank location, volume and minimum operational level as decision variables and assumes that tanks are cylindrical with a predetermined height-diameter ratio; that the emergency volume is a constant percentage of the tank water volume; and that the free-surface height is a constant percentage of the total water height. It also omits the riser based on its low impact on the objective function, determines maximum operational level based on the minimum operational level, volume, and height-diameter ratio; combines location and volume into a single decision variable; and simulates the tank's minimum operating level by another integer decision variable. Extended period simulations are performed using steady-state hydraulic analyses results for each time step. Tank water levels are adjusted according to the inflow/outflow conditions from the previous time step.

Vamvakeridou-Lyroudia *et al.* (2007) later conducted simulations to optimize tank placement in WDN design using a multiobjective approach (minimum cost and maximum benefit) where tank costs are given as a function of tank capacity. In their study, two decision variables were adopted for tanks: volume and normal minimum level. Other parameters were derived from these two.

The emergency volume was added manually at the end of the optimization process. These investigations were intended to reduce the gap between mathematical and engineering approaches to WDN tank design.

2.2.4.3 Resolution methods

Once the model is formulated and the tank decision variables selected, a number of different methods (linear programming, nonlinear dynamics, and evolutionary algorithms) can be employed in problem resolution. The optimization algorithm is coupled with a simulation model to solve equations of continuity and conservation of energy at each iteration. Resolution therefore requires both the optimization algorithm and simulation model.

Tank variables complicate WDN optimization by significantly increasing the total number of variables. Whereas tanks are associated with about 10 variables, the only variable associated with pipes is diameter. And each additional tank increases the number of possible solutions exponentially.

The introduction of genetic algorithms in WDN optimization makes it easier to optimize more complex systems (Savic et al. 2008). Murphy *et al.* (1994) used a standard genetic algorithm incorporating tank design to work out a very good solution to the “Anytown” network modeling problem. A more sophisticated genetic algorithm named SMGA (Structured Messy Genetic Algorithm) was used by Walters *et al.* (1999). Vamvakeridou-Lyroudia *et al.* (2005) used a fuzzy optimization model that proved equally effective.

Prasad (2007) added better constraint handling process to the methodology for tank design in the steady state genetic algorithm he used on the “Anytown” network design model. Other genetic algorithms such as NSGAI and SPEA2 have also been studied (Farmani et al. 2006).

Table 2.3 summarizes the optimization methods application to tank design in WDN.

Among the simulation models available for WDN design are AQUAGEO and EPANET (Brière, 2007). The latter is a well-known robust hydraulic solver tested for WDN (Reca & Martínez, 2006). Many other similar models are also available.

Table 2.3: Optimization methods application to Rank Design in WDN

Research paper	Tank decision variables	Formulation	Resolution method
Alperovits and Shamir (1977)	• Elevation	Single-objective	Linear programming (LP)
Murphy et al. (1994)	• Location	Single-objective	Genetic algorithm (GA)
	• Volume		
	• Minimum operational level		
Walters et al. (1999)	• Location	Multiobjective	Structured messy GA (SMGA)
	• Volume		
	• Minimum/maximum operational level		
Farmani et al. (2004)	• Location	Multiobjective	Nondominated sorting GA (NSGA)
	• Minimum/maximum operational level Diameter		
	• Bottom elevation		
Ostfeld (2005)	• Volume	Single-objective	GA
Farmani et al. (2005)	• Location	Multiobjective	NSGA
	• Minimum/maximum operational level		
	• Diameter		
	• Bottom elevation		
Vamvakieridou-Lyroudia et al. (2005)	• Location	Multiobjective	GA/fuzzy reasoning
	• Volume		
	• Minimum operational level		
Farmani et al. (2006)	• Location	Multiobjective	NSGA
	• Minimum/maximum operational level		
	• Diameter		
	• Height of water for emergency		
Vamvakieridou-Lyroudia et al. (2007)	• Volume	Multiobjective	GA/fuzzy reasoning
	• Minimum operational level		
Prasad (2007)	• Location	Single-objective	GA
	• Volume		
	• Minimum operational level		
	• Diameter-height ratio		
	• Emergency volume-total volume ratio		
Prasad (2010)	• Location	Single-objective	GA
	• Volume		
	• Minimum operational level		
	• Diameter-height ratio		
	• Emergency volume-total volume ratio		

2.2.5 Evaluation of both practices

2.2.5.1 Closing the gap between engineering practice and mathematical optimization models

Despite recent advances in mathematical models for WDN tank optimization, in engineering practice, a new tank is introduced into a WDN at a predetermined location and elevation (manual selection – step 1), followed by pipe optimization (mathematical model – step 2). The tank volume is decided through extended-period simulation (manual check – step 3) (Walski et al. 1987). This approach allows for only a limited number of trials on tank location, elevation and

storage prior to selection. The WDN optimization models used in step 2 include tanks and storage as decision variables, but often simplify assumptions on the simulation, calculation and grading of tank performance (Vamvakieridou-Lyroudia et al. 2007). Some of these models do not take tank shape into consideration and may produce either extremely high or flat tanks. In engineering practice, the use of optimization models to determine tank size is almost nonexistent (Walski, 2001). Where network optimization is even considered, engineers still prefer the traditional three-step procedure mentioned above, creating an important gap between engineering and mathematical approaches to WDN design optimization. Researchers are excited by the concept of optimization, but often do not work with practical engineering problems. Developers of optimization models are usually more experienced in mathematics and computer programming than they are in engineering practice, so that subtle nuances in the impact of various mathematical assumptions can easily escape them. Researchers and developers also often fail to understand the decision-making process involved in the design of WDN in real-world contexts. The literature abounds with models that fail to reflect the particular context in which engineers and planners make water-distribution decisions (Walski, 1995).

Another factor that could explain such a gap is the stochastic nature of networks themselves. In practice, WDNs do not display deterministic behavior, but are subjected to continuously varying conditions that have both deterministic and probabilistic components. A more realistic view of a network's performance can thus be obtained through a stochastic analysis that includes both deterministic and probabilistic parameters (Aly & Wanakule, 2004; Homwongs et al. 1994; Kretzmann & Van Zyl, 2004). Because of the nature of the stochastic problem, optimization cannot be achieved without forecast planning. On the other hand, large quantities of water cannot be stored given the random nature of network behavior and the risk of creating other problems (e.g. water quality). Therefore, in the optimization process of tanks in WDN, stochastic analysis can provide insights not seen with a deterministic analysis.

Stochastic analysis has been applied to water distribution systems in a number of studies (Damelin et al. 1972; Shu-li et al. 1996; van Zyl et al. 2012; van Zyl et al. 2008b; Wagner et al. 1988), though many of these focused mainly on pipe failures. van Zyl *et al.* (2008b) found that water demand has the greatest impact on tank size requirements. Nel and Haarhoff (1996) proposed a stochastic analysis technique for evaluating tank reliability, which employs a Monte

Carlo simulation of a system using stochastic models for user demand, fire demand and pipe failures. Van Zyl *et al* (2008a) refined this technique and proposed that tank design integrate failure risk at the most critical time in the year on a 10-year period – typically during peak seasonal demand.

2.2.5.2 Strengths and weaknesses of approaches to tank design in WDN

Given the issues around real-world application of optimization models discussed above, encouraging their use requires a clear examination of the strengths and weaknesses of different approaches to tank design in WDN and a process for selecting the approach most appropriate to a given WDN project from the limited variety of models currently available. Table 2.4 summarizes the strengths and weaknesses of each approach.

Table 2.4: Strengths and weaknesses of different approaches

Model / Approach	Strengths	Weaknesses
Guidelines	<ul style="list-style-type: none"> ▪ Simplicity ▪ Straightforward implementation ▪ Quick estimation 	<ul style="list-style-type: none"> ▪ Rigidity and inconsistency ▪ Do not provide the default risk ▪ Oversizing ▪ Focus only on tank volume
Single-objective	<ul style="list-style-type: none"> ▪ Simplicity and accessibility ▪ Rapid indication of the problem's nature 	<ul style="list-style-type: none"> ▪ No alternatives ▪ Not suitable for WDN
Multiobjective	<ul style="list-style-type: none"> ▪ Flexibility ▪ Very suitable for WDN ▪ Realistic 	<ul style="list-style-type: none"> ▪ Difficulty of implementation

2.2.5.2.1 Guidelines-based approach

Guideline-based models are straightforward to implement and require less data collection. Guidelines do not completely specify the design of tanks, leaving a great deal of room for engineering judgment. Where guidelines exist, they usually focus on volume and not on factors such as bottom and top elevation and location. Such approaches are useful to provide

designers a quick estimate of cost early on in the planning process. However, guidelines can be inconsistent between jurisdictions, with tank capacity requirements ranging from less than 50% to 200% of the maximum daily production of the water treatment plant (IWSA, 1955). Intrinsic safety factors associated with guidelines are unknown (Nel & Haarhoff, 1996). There is a risk of oversizing the tank, even when the WDN could operate efficiently with less storage capacity (Piqueiro et al. 1991). Guidelines have to consider a wide range of possible conditions and are thus likely to overstate tank size. Reliance on guidelines may result in the use of large tanks and the quality of the distributed water can become problematic as water sits in the tank for long periods of time.

2.2.5.2.2 Single-objective optimization models

Single-objective models are simple and accessible, and a wide range of tools is available to implement these models. This type of optimization provides policymakers with much needed information on the nature of the design problem (Savic, 2002). Single-objective models are flexible, but do not provide a set of alternatives representing a compromise between different objectives. Neither do they reflect the impact of cost increases or allow for comparison with alternative means of improving reliability (Walski, 2001). Tank size is considered only in an effort to minimize the overall network cost and issues such as water quality are not addressed. Single-objective models are not well adapted to WDN optimization (De Neufville et al. 1971). WDN with storage tanks should be designed for reliability and recognize the important contribution of tanks in achieving this goal. Cost minimization is unsuitable as an overriding objective because uncertainty about future demand and cost variations render long-term projections impossible. Although several tank variables can be included in these models, they do not all contribute to the calculation of cost.

2.2.5.2.3 Multiobjective optimization models

Multiobjective models offer a range of solutions, and it is up to the designers to select the one that best meets their priorities and criteria. These models provide a set of alternatives that enable compromise between different objectives and can take into account both network reliability and long-term resilience (Cohon, 2004). Multiobjective models are well adapted to the WDN design problem as they handle the issue of data randomness better (the random nature of certain

variables such as consumer demand, pipe roughness, tank levels and supply source levels) by maximizing system reliability and water quality in tanks. Multiobjective models have the potential to account for more of the factors considered in real decision-making if the objectives can be formulated accurately. The consideration of multiple objectives also promotes more appropriate roles for the different parties involved in decision-making and planning around WDN: the analyst generates alternatives for water distribution executives to consider in their selection of the most suitable design (Cohon, 2004). Tank variables included in these models contribute to the different objectives such as cost, reliability and water quality.

2.2.6 Choosing the appropriate model

The choice of model should consider these strengths and weaknesses, but must also take into account other aspects of a project. A step-wise decision process is proposed here to facilitate the selection of a project-appropriate model for optimizing WDN.

The first question in optimizing WDN that include tanks is whether or not a new tank is needed. In some cases, when a network is already in place and has to be extended to meet increasing demand, the introduction of a new tank can be useful to alleviate the burden on existing tanks, and thus avoid the use of large pipes for emergency response. Hydraulic simulations of future water demand scenarios are required to assess the need for a new tank. Capacity is most likely to fail during peak demand or in conditions where peak demand is combined with fire flow, making it essential to study the network's behavior under peak demand scenarios in order to evaluate the need for new storage. If a tank is not required, the literature recommends a number of approaches related to WDN pipe optimization (Savic and Walters 1997; da Conceicao Cunha and Sousa 1999; Doby et al. 2004; Prasad 2004; Ekinici and Konak 2009; Guistolisi et al. 2009). After the need for a new tank is established, its optimal design must be determined. The following step-wise procedure allows for a reasonable selection (see Figure 2.2):

1. Select tank decision variables. This choice depends on technical constraints such as availability, topography, aesthetics, pressure gradients, problem size, etc. In some cases, when constraints are imposed by property ownership, zoning, proximity to transmission mains, etc., the number of possible sites for any new tank is usually limited to one or two

and optimization is not really required to compare them. In such cases, tank location is not considered as a variable.

2. Some variables will be covered by local guidelines and these should be examined to see if there is room for optimization. Guidelines usually focus only on volume, leaving room for optimization on other variables. Review local guidelines on WDN tank design and compile a complete history in order to assess the reliability of guidelines. This involves evaluating the historical use of guidelines. If a tank designed in accordance to guidelines often runs dry, fails to operate correctly or presents serious quality problems, the guideline's reliability is questionable. Find out how flexible the enforcement of guidelines is regarding optimization strategies that differ from these guidelines;
3. Based on the results of this investigation, the analyst can then choose to: (1) prioritize guidelines for the variables they cover and combine these with engineering analysis, which may include optimization, for other variables; (2) prioritize optimization for all variables but include some guideline constraints. Where guidelines and engineering judgment (without optimization methods) are involved in tank design, any of the pipe-focused optimization models can be used.
4. Where an optimization model will be used, conduct a preliminary analysis to determine the most relevant objective for network design (capital costs, meeting future demand, hydraulic performance, reliability, etc.).
5. Select either the single-objective or multiobjective model, depending on the most relevant objective. For instance, if the most relevant objective is to minimize the capital costs of meeting future peak demand under minimum pressure, the single-objective formulation will prove most appropriate (Filion 2008). Select the model sub-type that can best solve the optimization problem. This choice depends on the problem size. For large problems, evolutionary algorithms are more efficient.
6. Compare the results with those of other model types (if available) or to guideline recommendations.

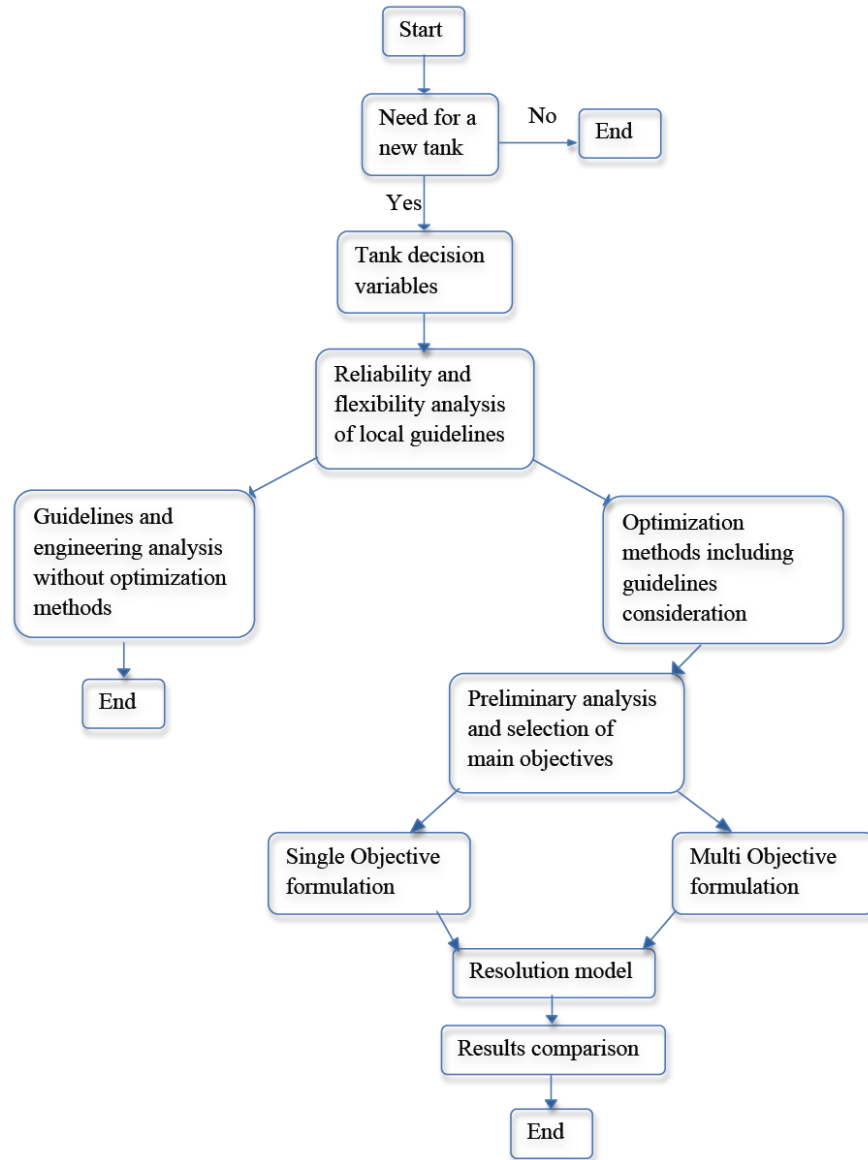


Figure 2.2: Procedure for selecting appropriate approach

A case study using this procedure would reveal how local constraints influence decisions and how this methodology could be adapted and improved. In fact, the choice is not binary but rather involves prioritizing a spectrum of guidelines and optimization approaches. Since most systems have multiple pressure zones, the application of these ideas should consider the issues that arise in these cases. In fact, in multiple pressure zones, the source is usually located in one of the lowest zones and the system is fed through pumps. In such cases, it is best to place one tank in each zone so that each zone can operate almost independently. In multiple pressure-zone systems

where the source is located in the highest zone and lower zones are fed through pressure-reducing valves, most of the storage should be kept in the higher zones. In that way, stored water can feed virtually any zone (Walski, 2000).

2.2.7 Discussion

Multiobjective models have been used to include tank design in WDN optimization (Farmani et al. 2006; Vamvakeridou-Lyroudia et al. 2007). An increasing number of professionals and researchers believe multiobjective design can lead to much more robust WDN, albeit at slightly higher cost. The optimization model developed by Farmani *et al.* (2005) was the first to size and locate all system components, including pipes, pumps and tanks (Filion 2008). In the past decade, WDN optimization models have focused on pipes (Ekinici & Konak, 2009), tank operation (Fang et al. 2010) or the impact of tank size on energy consumption (Ghimire & Barkdoll, 2010). Models that specifically seek to optimize tank size in WDN are still quite rare.

In terms of efficiency, genetic algorithms have proven strengths (Savic et al. 2008) and the number of recent articles drawing on these algorithms is an indication of their preferred status in the field. However, more research is needed to improve the resolution time of these algorithms. Multiobjective models respond better to the desire to choose between different alternatives, but there is a need for better guidance when it comes to pinpointing the optimal solution.

2.2.8 Conclusion and recommendations

Analysis of the various models of WDN optimization reveals that although many still strongly focus on pipe optimization and tank operation, a handful of more recent models attempt to integrate tank design. These efforts should be encouraged and steered toward greater applicability in real-world practice. Storage tanks have a valuable role to play in WDN and their optimal design contributes importantly to network reliability and efficiency. It is however evident that some variables continue to be ignored or inaccurately assumed constant in these models. Future models should include real-world network problems and stochastic considerations to reduce the gap between optimization models and engineering practice.

CHAPITRE 3 ARTICLE 2: MULTI-OBJECTIVE DESIGN AND DECISION-MAKING PROCESS IN WATER DISTRIBUTION SYSTEMS

3.1 Présentation de l'article

Dans cette section, l'objectif de l'article sera présenté, ainsi qu'une conclusion sur ce travail dans le cadre de la thèse. Cet article présente une approche multiobjective de conception des réseaux de distribution d'eau potable (RDEP) incluant les réservoirs dans un environnement déterministe. L'approche contient des paramètres à personnaliser en fonction du problème traité. Elle peut être utilisée pour la conception optimale d'un nouveau réseau ou l'extension d'un réseau existant, moyennant une intervention dans le code source du programme développé. L'approche propose les étapes à suivre pour un problème donné de conception d'un RDEP et constitue un bon outil d'aide à la décision.

3.1.1 Objectif

Le mot «optimisation» à lui seul n'a aucun sens car les résultats de cette optimisation dépendent des objectifs, des contraintes et des données utilisées. Le seul objectif de minimiser le coût devient de plus en plus une approche inacceptable dans la conception des RDEP, d'autant plus que l'introduction des réservoirs s'accompagne souvent de problèmes liés à la qualité de l'eau. Les investigations de la revue de littérature de l'article 1 ont permis entre autre de prendre en compte les incertitudes liées aux données lors de la conception des RDEP. L'approche proposée dans cet article, bien qu'implémentée dans un environnement déterministe, peut intégrer implicitement ces incertitudes en ajoutant un paramètre de fiabilité comme partie intégrante des objectifs de conception. Prenant en compte le coût, la qualité de l'eau et un indicateur de fiabilité du réseau comme objectifs simultanés de l'optimisation du RDEP, cet article donne les étapes à suivre depuis la formulation jusqu'à la résolution en plus de proposer des équations à paramètres personnalisables. L'applicabilité de l'approche est testée sur un réseau complexe et réaliste ayant fait l'objet de compétition lors d'une récente conférence en Australie.

3.1.2 Conclusion

Il est observé que l'approche proposée, bien que n'étant pas exhaustive, permet de s'attaquer efficacement à des réseaux réels et complexes en réduisant le nombre de variables éventuelles ainsi que leurs intervalles de variation. L'approche de formulation multiobjective offre beaucoup plus de compromis entre les divers objectifs et aussi plus d'alternatives aux gestionnaires des RDEP. La prise en compte de la demande aux nœuds dans l'évaluation de la qualité de l'eau pendant le processus d'optimisation permet de minimiser l'impact sur les consommateurs d'une qualité d'eau médiocre. Par ailleurs, la considération des émissions de gaz à effet de serre comme troisième objectif (en plus du coût et de la qualité de l'eau) intègre l'aspect environnemental dans la conception favorisant ainsi un développement durable.

Cet article est soumis à "Engineering Optimization".

Batchabani, E. and Fuamba, M. (2013b). "Multi-objective design and decision-making process in water distribution systems." Engineering Optimization, Submitted.

3.2 Multi-objective design and decision-making process in water distribution systems

Essoyeke BATCHABANI, Ph.D. Candidate, Department of Civil, Geological and Mining Engineering, École Polytechnique de Montréal, P.O. Box 6079, Station CV, Montreal, Quebec, Canada, H3C 3A7. Email: essoyeke.batchabani@polymtl.ca

Musandji FUAMBA, P.Eng. Ph.D., Professor, Department of Civil, Geological and Mining Engineering, École Polytechnique de Montréal, P.O. Box 6079, Station CV, Montreal, Quebec, Canada, H3C 3A7. Email: musandji.fuamba@polymtl.ca

3.2.1 Abstract

Water Distribution Systems (WDSs) design and operation are usually done on a case-by-case basis. Numerous models have been proposed in the literature to solve specific problems in this field. The implementation of these models to any real-world WDS optimization problem is left to the discretion of designers who lack the necessary tools that will guide them in the decision-making process for a given WDS project. In order to support a decision process in this field, the present article suggests a step-by-step approach to solve the multi-objective design problem. A genetic algorithm (GA) is proposed as the optimization tool and the targeted objectives are: (1) to minimize the total cost (capital and operation), (2) to minimize the residence time of the water within the system and (3) to maximize a network reliability metric. The results of the case study show that preliminary analysis can significantly reduce decision variables and computational burden.

Keywords: decision-making, genetic algorithm, multi-objective, optimization, water distribution systems.

3.2.2 Introduction

Water Supply involves intricate systems and components such as treatment plants, distribution systems, pumps and tanks. Water Distribution Systems (WDSs) are designed to meet actual and future water demand under a set of constraints such as minimum pressure and good water quality. However the factors associated with the construction and maintenance of WDSs are

complicated and expensive. In the actual context of highly increasing demand, cost effective design improvement alternatives need to be identified and analyzed.

Numerous WDS optimization models including complex algorithms have been proposed over the years. Such models include linear programming (Alperovits and Shamir 1977; Quindry et al. 1979; Fujiwara 1990; Sherali et al. 1998; Samani and Mottaghi 2006), nonlinear programming (Djebedjian et al. 2000 ; Lansey & Mays, 1989), and dynamic programming (Liang, 1971). In recent years, a number of evolutionary algorithms (EAs) have been introduced for optimizing the design and operation of WDSs. These EAs include genetic algorithms (Dandy et al. 1996; Simpson et al. 1994), ant colony optimization (Maier et al. 2003), the shuffled leaping frog algorithm (Eusuff & Lansey, 2003), particle swarm optimization (Suribabu and Neelakantan 2006; Montalvo et al. 2008a; Montalvo et al. 2008b), harmony search (Geem, 2006), genetic heritage evolution by stochastic transmission (Bolognesi et al. 2010) and differential evolution (Suribabu 2010; Vasan and Simonovic 2010). Among all these EAs, the Genetic Algorithm (GA) represents the oldest and most established evolutionary algorithm that has been applied to WDSs. As it stands, plenty of options and variants of GAs have been developed: different types of selection, crossover and mutation can be found in the literature, different coding for the variables are possible, and different hybrid versions of GAs have been proposed. A recent study by Marchi *et al.* (2014) comparing EAs for the optimization of WDSs showed the robustness of the GA compared to the other techniques in that the results obtained are less sensitive to the parameters.

Many WDS optimization models submitted in the literature are developed on a case-by-case basis, making it hard for designers to implement them on other real world WDS design problems. For example, some models tailored for Anytown network (Walters et al. 1999; Farmani et al. 2004; Farmani et al. 2005; Olivera Sousa et al. 2005; Farmani et al. 2006; Prasad 2007; Vamvakeridou-Lyroudia et al. 2007; Prasad 2010), Hanoi network (Cembrowicz et al. 1996; Perelman et al. 2009), or New York Tunnels network (Khu & Keedwell, 2005) that implement a chosen network does not always represent the different scenarios faced in the real life. Any changes in the network would then lead to a new approach to be considered as no general means are available when dealing with common points that can be successfully applied to several networks. One implication to this observation is a lack of necessary tools guiding the decision-making process for any given random WDS project. Designers need documents that can

help them to map out paths at least on where they can start and to provide them with some actionable insights that can be adaptable to various networks. The issues of the need for a tank, the number of tanks, the need for a pump or the number of pumps have to be addressed where applicable. This article aims to contribute to this effort by suggesting helpful insights in a step-by-step approach to deal with WDS design problems in real practice. The issues raised above are discussed throughout the approach and the suggestions provided here for WDS design are intended to formulate and solve an optimization problem. The aspects emphasized in this article are: (1) the demand for a preliminary analysis before the optimization step, (2) the requirement for reducing as much as possible the number and the range of decision variables, and (3) the need for generalizing the approach as much as possible. The methodology uses a genetic algorithm to solve the formulated optimization problem. The algorithm coded in C++ is coupled with the well known simulation model *EPANET* (Rossman, 1999). The design options and criteria discussed in this article are not comprehensive but engineers frequently encounter them in real practice.

The next section of this article describes the proposed approach which a WDS designer could adopt to deal with the optimization problem. Results from the application of the approach on a real world case study are presented in section 3. Finally, section 4 offers a conclusion and a recommendation for future applications.

3.2.3 Approach

This article suggests practical steps to follow in order to achieve a given WDS design optimization problem from the formulation to the resolution. This section begins by identifying and formulating the main variables associated with the WDS design, and the section right after defines the performance criteria which may be used in the optimization process. The following Section then introduces the method proposed in this article to solve the formulated multi-objective optimization problem. Figure 3.1 gives an overview of the three steps.

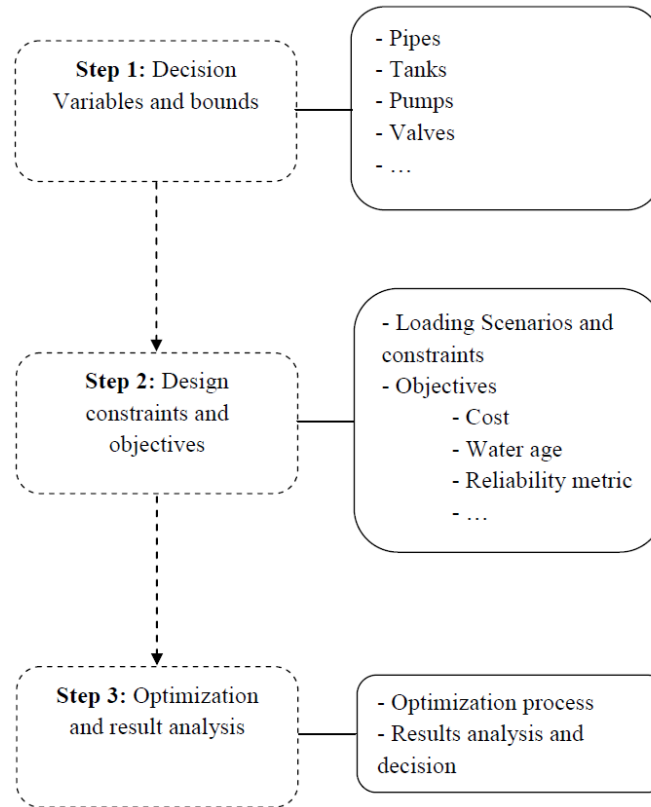


Figure 3.1 : Steps overview of the proposed approach

3.2.3.1 Identify and formulate the design options and bounds

The objective of this section is to identify and formulate the main variables usually associated with the WDS design. Four network components are considered to be the main sources of the design variables. Each of these components with the associated variables is described in the following subsections.

3.2.3.1.1 Pipes

The first variable of each pipe could be the design option. Several options may be considered depending on the problem nature (new design, rehabilitation or network expansion). These options include: (1) cleaning, (2) lining, (3) duplicating, (4) replacing, or (5) doing nothing. The “option” variable can be modeled as an integer, ranging from 1 up to the number of available options. For a new design problem, this variable can be ignored.

The second variable for each pipe is the diameter size to be used. This diameter has to be chosen in a list of discrete commercial diameters. Although this list may be extensive, it is important to

reduce the variable range in order to guide the optimization process toward a reasonable solution space. In this approach, a diameter is computed to be the upper limit for the diameter size. This diameter (D_{sup}) is based on a maximum velocity V_{max} imposed by the designer. The maximum water demand Q_{max} for a single time step Δt has to be computed based on the loading scenarios data (see section 2.2.1). The upper limit diameter (D_{sup}) is then given in Equation (3.1) where Q_{max} and V_{max} are expressed in SI. D_{sup} can be interpreted as the minimum diameter of a pipe that would carry Q_{max} at a maximum velocity of V_{max} .

$$D_{sup(mm)} = \text{Nearest Upper Commercial Diameter} \left(1000 * \sqrt{\frac{4*Q_{max}}{\pi*V_{max}}} \right) \quad (3.1)$$

The commercial diameters have to be sorted in ascending order with a rank assigned to each size. If RD_{sup} denotes the rank assigned to D_{sup} , the “diameter size” variable can be modeled as an integer, ranging from 1 to RD_{sup} . For a network rehabilitation problem where the pipe sizes cannot be modified, this variable can be ignored.

3.2.3.1.2 Tanks

There are several variables which could be theoretically associated with tanks in WDS design optimization (Batchabani & Fuamba, 2014b). In practice, these variables are highly case specific. With this approach, some parameters are computed in order to guide a decision regarding the usage of tanks.

If the number of tanks is unknown, the approach suggests computing the number (N_{zones}) of pressure zones. According to Walski (2000), pressure zone hydraulic grade lines should differ by roughly 30.48 m from one pressure zone to the next. Based on this assumption, N_{zones} is given by Equation (3.2) where E_{min} and E_{max} are respectively the minimum and maximum node elevations of the network. The suggestion made is to provide one tank in each pressure zone, giving a number of potential tanks equals to N_{zones} .

$$N_{zones} = \text{Round} \left(\frac{E_{max} - E_{min}}{30.48} \right) \quad (3.2)$$

If the tank bottom elevation variable has not been predetermined, the designer needs to set the bounds on this variable for the optimization process in order to reduce the search space. The recommended bounds are provided in each pressure zone in Equations (3.3) and (3.4) based on the elevation of the pressure zone that the tank will serve. In these equations, $A_{min,zone_i}$ and

$A_{max,zone_i}$ are respectively the minimum and maximum tank bottom elevation in the pressure zone i ; $E_{min,zone_i}$ and $E_{max,zone_i}$ the minimum and maximum node elevations of the pressure zone i ; h_{min} with h_{max} the minimum and maximum allowable pressures. These bounds are computed so that the tank could give enough pressure to the highest customer and non-excessive pressure to the lowest one. The “tank bottom elevation” can be modeled as a real variable, ranging from $A_{min,zone_i}$ to $A_{max,zone_i}$ in each pressure zone.

$$A_{min,zone_i} = E_{max,zone_i} + h_{min} \quad (3.3)$$

$$A_{max,zone_i} = E_{min,zone_i} + h_{max} \quad (3.4)$$

Tank volume is relegated by the capacity of components such as reserves for production, emergency, fire or balancing depending on the use of the storage. A tank should basically supply the balancing reserve for useful storage. Emergency, production and fire reserves can be estimated based on local guidelines. This article offers an approach to estimate the balancing storage $B_s(m^3)$ before moving forward to the optimization step. The estimation for this reserve is based on a full time pumping during the entire extended-period simulation time. The approach has to be applied on the whole network and on the independent pressure zones in order to evaluate the balancing reserve for different tanks. The steps to follow are: (1) calculate the average water demand $Q_{moy}(m^3/period)$ which represents the uniform pumping rate along the extended period simulation; (2) compute the tank storage volume for each time period i , $S_i(m^3/period)$ by utilizing the continuity principle as in Equation (3.5) with an initial storage S_0 of zero; (3) the balancing storage is then given by Equation (3.6). Q_i in Equation (3.5) denotes the total water demand ($m^3/period$) in the period i for the considered pressure zone.

$$S_i = Q_i - Q_{moy} + S_{i-1} \quad (3.5)$$

$$B_s = \max_i S_i + |\min_i S_i| \quad (3.6)$$

Evaluating the balancing reserve coupled with the uniform pumping rate is useful especially for defining if the existing pumps and tanks are sufficient or in need of upgrading. It is also important for the optimization process to reduce the number and the range of decision variables. With the estimated needed volume known, an assumption has to be made on the shape (diameter-height ratio) to set the tank diameter.

The choice in determining tank location in real-world problems can be constrained by the availability of land already owned by the water utility or by the urban texture. In order to define this potential variable, specific information is required. If several nodes are regarded potential points for a tank location in a given pressure zone, this variable for the corresponding tank can be set as an integer starting from 1 up to the set number of candidates.

3.2.3.1.3 Pumps

There are two types of variables for pumps: pump design variables and pump operation variables. For pump design, the two parameters involved are the flow and the head. The minimum required flow is Q_{moy} (in Equation (3.5)) and the minimum required head $H_{p,min}$ (if head losses are not assumed) is given in Equation (3.7) where E_{source} signifies the elevation of the pumping source. The couple $(Q_{moy}, H_{p,min})$ has to be regarded in selecting pump curves for each pressure zone.

$$H_{p,min} = E_{max} + h_{min} - E_{source} \quad (3.7)$$

For the available pump curves, it is important to compute the maximum power with the corresponding flow and head to ascertain if these curves meet the network needs or if it is necessary to combine them for this purpose. For a rehabilitation problem, these calculations help to figure out if actual pumps need to be upgraded or not.

Pump operation can be either time-based or based on threshold tank levels or both. If only time-based controls are chosen, the pump operation variables can be formulated as binary numbers corresponding to pump status (ON/OFF) in each time step. For the threshold tank levels based operation, the variables are the tank levels leading to pump switching (ON/OFF).

3.2.3.1.4 Valves

Different types of valves can be used to limit the pressure or flow at a specific point in the network. These types include pressure reducing, pressure sustaining, pressure breaker, flow control or general control valves. If valves are to be included, the decision variables could be the location and the setting. Typically, the location is imposed by the configuration of the network, leaving the setting as the only adjustable variable. For a rehabilitation problem, the existing setting may still be effective and no change would then be required.

3.2.3.2 Define and formulate the design performance evaluation criteria

In order to evaluate the design solutions characterized by assigning values to the variables discussed in the previous section, the performance assessment metrics have to be defined. Only feasible solutions are normally desirable and among these solutions, the best ones are found in the optimization process. This section begins with discussing the design constraints to define feasible solutions before dealing with design objectives which guide the optimization process.

3.2.3.2.1 Loading scenarios and design constraints

Loading scenarios with the corresponding design constraints, set by the designer, may include normal and abnormal operation conditions. For these scenarios to run effectively, the simulation must be balanced by a specified period that can be timed in an hourly, daily, weekly or any other balancing period. The hydraulic and water quality time-steps have to be selected based on the desired level of precision and the resulting computation burden. The design constraints can be measured in terms of minimum and maximum pressures provisioned at demand nodes for each time step within the design balancing period. Depending on the designer requirements, it is possible for many other constraints to develop. Handling constraints compels that a penalty function is used, so that infeasible solutions are penalized depending on their level of infeasibility, while feasible solutions come up with a penalty score of zero. Each constraint violation that appears is reported in this penalty function.

3.2.3.2.2 Design objectives

Design objectives are the functions that can be maximized or minimized in the optimization process. WDSs are normally built and operated so as to: minimize capital and operation cost, maximize water quality, and maximize various system reliability metrics that include environmental considerations. Based on the designer's criteria, only one objective or a combination of several objectives may be at play. Some potential common objectives are discussed in the next subsections.

Costs: the total cost includes capital costs and operational costs. Capital costs consist of component costs of pipes, tanks, pumps, valves and other possible problem specific components.

The operational costs, on the other hand, are basically associated with the system power usage during the design horizon; however, other problem specific operation costs can be present. The costs must take into account the lifetime of each component and the discount rate. The pipe costs are function to the pipe diameter and length while tank costs basically subject to the tank volume. The pump investment costs are representative to pump power and the valve costs are inherent on their diameters. These costs are problem specific and part of input data.

Water Quality: several parameters can be used to characterize the water quality: microbial growth, residual chlorine decay or hydraulic residence time. In numerical modeling, most of water quality models use the hydraulic residence time to compute different substance concentrations. This residence time which is expressed in the form of water age is used as water quality metric in the optimization process that can be defined in several ways. To account for water age impact on consumers, water age would be measured at non-zero demand nodes only while giving more weight to nodes with larger demands. The water age measure in Equation (3.8) includes these computations and can then be adopted in the optimization process.

$$WA_{net} = \frac{\sum_{i=1}^{N_{junc}} \sum_{j=1}^{N_{time}} k_{ij} * Q_{dem,ij} * WA_{ij}}{\sum_{i=1}^{N_{junc}} \sum_{j=1}^{N_{time}} Q_{dem,ij}} \quad (3.8)$$

In Equation (3.8), WA_{net} refers to the network water age, WA_{ij} tells the water age at junction i at simulation time t_j , k_{ij} is a binary variable defined as 1 if WA_{ij} is greater than a threshold WA_{th} and zero otherwise, $Q_{dem,ij}$ points to the demand at junction i at time t_j , N_{junc} is the number of system junctions and N_{time} counts the number of simulation time steps.

Reliability: various metrics can be used to describe the network reliability. It is up to the designer to choose the most appropriate reliability metric for his specific problem. Three reliability measures are presented in this article: the network resilience, the greenhouse gas emissions and the leakage.

The Network Resilience I_n was introduced by Prasad (2004) to improve the earlier resilience index I_r introduced by Todini (2000) as a way of measuring WDS reliability to account for means taken to mitigate the risk of failures. The Network Resilience incorporates the effects of

both surplus power and reliable loops thus demonstrating the effect of redundancy. This index is computed for each time step i as in Equation (3.9) where np numbers the amount of pipes connected to a node, D shows the pipe diameter, h reveals the node pressure, N_r totals the number of tanks, Q_r presents the tank outflow, H_r concerns the total head of the tank, N_b sums up the number of reservoirs, Q_b expresses the reservoir outflow, H_b speaks to the total head of the reservoir, N_p displays the number of pumps Q_p and H_p are respectively the rated discharge and the head of the pump.

$$In_i = \frac{\sum_{j=1}^{N_{junc}} \left(\frac{\sum_{k=1}^{np_j} D_k}{np_j * \max\{D_k\}} * Q_{dem,ij} * h_{i,j} \right) - \sum_{j=1}^{N_{junc}} \left(\frac{\sum_{k=1}^{np_j} D_k}{np_j * \max\{D_k\}} * Q_{dem,ij} * h_{min,j} \right)}{\left(\sum_{j=1}^{N_r} Q_{r,i,j} * H_{r,i,j} + \sum_{j=1}^{N_b} Q_{b,i,j} * H_{b,i,j} + \sum_{j=1}^{N_p} Q_{p,i,j} * H_{p,i,j} \right) - \sum_{j=1}^{N_{junc}} \left(\frac{\sum_{k=1}^{np_j} D_k}{np_j * \max\{D_k\}} * Q_{dem,ij} * h_{min,j} \right)} \quad (3.9)$$

The Greenhouse gas emissions are basically the result of the emissions associated with the energy required for manufacturing, transportation and installation of WDS components and the power usage from the operation of pumps. The designer may want to minimize these emissions for the sake of a sustainable development. The emission rates are part of the input data. A monetary value can be assigned to GHG emissions, enabling their inclusion in the total cost. However, Wu *et al.* (2010b) recommended from the results of their study to separate these objectives in the optimization of WDS. A GHG objective has already been considered in some previous studies for the optimization of WDS (Roshani et al. 2012; Wu et al. 2010b; Wu et al. 2010a).

Leakage in a WDS can be the result of mechanical failure from pipes or other components and from pipe junctions or pipe walls in aging networks. According to Tucciarelli et al.(1999), the background leakage in half-pipes linked to node i is calculated using Equation (3.10) where λ is a loss exponent, L the pipe length, θ the leak per unit surface of the pipe linking nodes i and j .

$$q_{i,t}^{leak} = (h_{i,t})^\lambda \sum_{j=1}^{np_i} \frac{\pi}{2} h D_{ij} \theta_{ij} L_{ij} \quad (3.10)$$

3.2.3.3 Solve the multi-objective optimization problem

The objective of this section is to offer the method proposed in this article to solve the multi-objective optimization problem formulated in the previous sections. The optimization and

simulation tools are introduced and then the optimization process is presented. Finally, the optimization results are to be analyzed so as to select a proper design.

3.2.3.3.1 *Genetic Algorithm*

Although many advanced Genetic Algorithms (GA) do exist in the literature and problem specific genetic operators could be developed, a steady state GA is used in this article and adapted for multi-objective optimization. For this purpose, a GA library called GALib (Wall, 1996) is employed which provides necessary tools for GA and its customization, along with its built-in operators. The following paragraphs discuss the selection of some GA parameters to effectively perform the optimization process. However, it is only indicative and other values can be assigned to these parameters depending on the computation time required to evaluate a single solution.

The population size has typical values between 50 and 1000 depending on the size of the solution space and the complexity of the adaptation evaluation (Ahn, 2006; Goldberg, 1987). It is a critical parameter for the proper functioning of evolutionary algorithms. If the population size is too small, then the algorithm is trapped in local areas of the search space (local optimum). However, if the population size is too large, then the algorithm will unnecessarily waste computing resources. Given these constraints, the population size is suggested to be set as follows: (1) if the total number of variables N_{var} is less than 50, the population size can be 50; (2) if N_{var} is between 50 and 1000, the population size can be N_{var} ; (3) if N_{var} is greater than 1000, the population size can be 1000. This approach of assigning values can be realized as a working hypothesis taking into account the number of variables and fitting typical values.

There are different forms of stopping criteria such as the number of generations, the scoring of the best solution or the convergence of the population which could be used in a GA process. In the coded program, the number of generations is considered as the stopping criterion and taken proportionally to the population size with a factor of 10. With this approach the number of generations falls in a range between 500 and 10,000 that seem logical as a good exploration setting of the solution space.

If the probability of mutation p_m is too high, the algorithm degenerates into a random search, which is undesirable since the algorithm must evolve towards the optimal solution. Typical values of the probability of mutation are between 0.001 and 0.01 depending on the problem

nature. Another method is to adopt a probability of mutation inversely proportional to the population size. In the proposed methodology, this probability can be set as follows: (1) if the population size is less than 100, p_m can be 0.01, (2) if the population size is greater than 100, p_m can be the inverse of the population size. The probability of mutation garnered by this parameter reaches between 0.001 and 0.01 as is commonly found. This event of assigning values can be seen as a working hypothesis taking note the population size and fitting typical values.

Typical values of the probability of crossover are rated between 0.5 and 1 depending on the case, although the probability of 0.9 is much used in the implementation of genetic algorithms. It is recommended that GA uses a moderate population size, a high probability of crossover and a low probability of mutation, even as the algorithm can be relatively insensitive to applied probability values. In this article, a value of 0.9 is proposed for the probability of crossover.

For the genome representation, the program uses a binary-to-decimal genome to handle the optimization variables. This genome is an implementation of the traditional method for converting binary strings to decimal values. To use this genome, what is required is the ability to create a mapping of bits to decimal values by specifying how many bits will be used to represent what bounded numbers.

3.2.3.3.2 *EPANET solver*

EPANET is executed in command mode. A series of DOS commands are introduced in the code to run the simulation and save results into a text file. The saved report is a comprehensive report and the results recorded at each time step for the different variables are: (1) demand nodes: actual demand, total head, pressure, quality; (2) reservoirs: net inflow, total head, pressure, quality; (3) tanks: net inflow, total head, pressure, quality; (4) pipes: flow, velocity, unit headloss, status; (5) pumps: flow, quality, head, status (6) valves: flow, velocity, headloss, status, *EPANET* will also record the energy consumption for the simulation period.

3.2.3.3.3 *Pareto ranking*

For a population (P) of solutions provided by the GA, a score has to be assigned to each single solution corresponding to its rank in (P). This rank reflects the degree of dominance of the solution in this population. The constrained optimization problem is reduced to an unconstrained

optimization problem reported in Equation (3.11). Note the equivalence between maximizing reliability and minimizing –reliability. The sign “-” is used in the case the reliability metric chosen is to be maximized. If the designer elects a metric which is to minimize, this sign is removed. The penalty value is added to each term of Equation (3.11).

$$\min F(x) = (f_1(x) = \text{Cost}, f_2(x) = \text{Water Age}, f_3(x) = -\text{Reliability}) \quad (3.11)$$

In Equation (3.11), x refers to the set of problem variables defined earlier. Each solution to the problem is associated with an objective vector to be minimized. Given two solutions to the problem x^1 and x^2 , x^1 is said to dominate x^2 if and only if $F(x^1)$ is partially less than $F(x^2)$; i.e., if and only if the conditions of Equation (3.12) are satisfied.

$$f_i(x^1) \leq f_i(x^2) \forall i = 1, 2, 3 \text{ and } \exists i \in \{1, 2, 3\} / f_i(x^1) < f_i(x^2) \quad (3.12)$$

A solution x^0 is said to be Pareto optimal or efficient with regard to a set of solutions if and only if there does not exist any solution in this set dominating x^0 . The set of non-dominated solutions forms the Pareto front. The latter is the one analyzed by the decision-maker to choose the solution it deems appropriate.

There are various Pareto ranking methods in a given population (Goldberg 1989; Fonseca and Fleming 1993; Belegundu et al. 1994). A comparative study by Mallor et al. (2003) showed that Fonseca & Fleming's method is stable and produces small errors in classification. This classification is the one used in this case to assign a rank to a solution x in P as shown in Equation (3.13). In this classification, the rank n of a solution x is the number of solutions that dominate x in P . The non-dominated solutions, therefore, all have the same rank of zero, while those seen as dominated solutions have ranks between 1 and $k - 1$, where k is the size of P .

$$\text{Rank}_{\text{Pareto}}(x) = \text{Number of solutions dominating } x \text{ in } P \quad (3.13)$$

For this purpose, a basic dominance function was created in the program. This function takes as arguments two vectors of three-dimensional real x^1 and x^2 . It returns a Boolean variable that is true when x^1 dominates x^2 according to Equation (3.12). Initializing the rank of x to 0, for each solution x^i in P , $i = 1$ to k , if the function dominance (x^i, x) returns true, the rank of x is incremented by one unit.

The score of each solution is set equal to its rank in the population. Also this score is to be minimized by GA. In this manner, the GA will favor non-dominated solutions in each

generation, allowing for fast convergence towards the optimal solutions. A population evaluator is then created in the program for this purpose. This function evaluates and sets the score of each genome in the current population during the GA evolving process.

3.2.3.3.4 Optimization process

The proposed program runs as follows (Figure 3.2):

- a. The program reads the input file: this allows recording relevant information on the network such as the number of pipes, pumps, tanks or demand nodes.
- b. From (a), the program computes indicative parameters: upper limit diameter, pressure zones, minimum and maximum tank bottom elevations which are later called if needed.
- c. The program provides the GA the form of a genome (solution) using the information recorded in (a) and (b) on the network configuration.
- d. The GA generates a set of random genomes in the same format that has been provided in (c). The number of genomes equals to the size of the population as computed earlier.
- e. The program retrieves each genome, converts its contents and writes the *EPANET*'s input file for simulation by the *EPANET* solver.
- f. The program uses the simulation results from (e) to calculate the values of the objective functions and the penalty associated with respect to constraints.
- g. The program computes the Pareto rank of each genome from (f). These ranks are retrieved by GA and measures the quality of each genome.
- h. The GA uses genome scores from (g) to generate a new improved population by selection, crossover and mutation.

Steps (e) to (h) are repeated until the stopping criterion is satisfied. The program saves the non-dominated genomes into a text file. Values of objective functions and penalty are recorded as well in the same file. These results are then analyzed by the decision-maker.

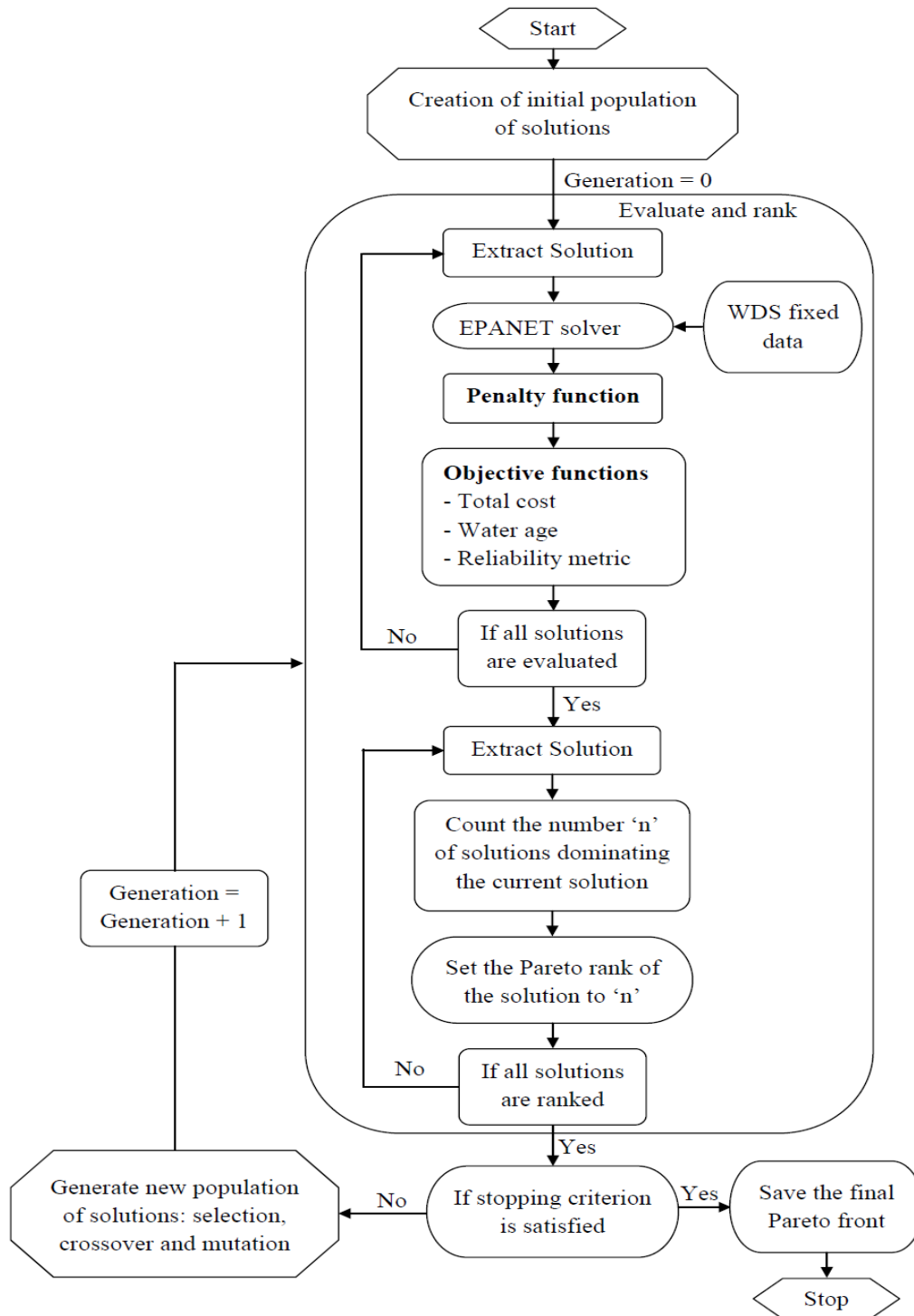


Figure 3.2 : Optimization process in the proposed approach

3.2.3.3.5 *Results analysis*

The adoption of one solution among the candidates from Pareto front is not universal and obvious. It hinges on the decision priorities since gaining on a goal deteriorates at least one other goal and vice versa which leads to compromise from the designer.

The decision-maker (DM) may set a maximum residence time and then would reject solutions with a residence time that exceed the set time. If DM is less concerned about water quality and more demanding on reliability or total cost, his options becomes more and more limited.

Pareto front is usually composed of a small number of solutions and the choice is relatively less polemical. If the difficult choice comes with a great number of non-dominated solutions, one of multiple decision methods such as AHP (Analytic hierarchy process) can be engaged. The Pareto approach for multi-objective optimization allows for making a reasonable choice in a small space of solutions even if it may not necessarily be the best one.

3.2.4 **Case study**

3.2.4.1 **Network description**

A large and complex network environment was addressed in the recent Battle of the Water Networks (Marchi et al. 2013) which was held at the Water Distribution System Analysis (WDSA) conference in Adelaide, South Australia.

The proposed approach this article brings is to apply the reduction of decision variables in this particular network environment as a means to increase decision effectiveness in solving network design issues. The full problem description pdf document and the associated EPANET input file can be obtained from the University of Exeter web page (<http://emps.exeter.ac.uk/engineering/research/cws/resources/benchmarks/expansion/d-town.php>). A summary description of the problem is outlined in the next paragraphs.

The municipality of D-Town is in need of a design project to cope with the increased water demand of the population and a general increase in the population size that has led to a new residential district in the Eastern part of the city. The calibration and simulation of the actual network model shows that the existing infrastructure is not able to meet the forecasted demand, and therefore an upgrade of the network is warranted. Figure 3.3 reports the layout of the D-

Town network which consists of five existing district metered areas (DMAs) requiring upgrades and the additional new zone to be designed. In total, The D-Town network consists of 399 junctions, 7 storage tanks, 443 pipes, 11 pumps and 5 valves, and a single reservoir. The optimization objectives contemplated by the water utility are to minimize capital and operational costs while minimizing greenhouse gas (GHG) emissions and improving water age.

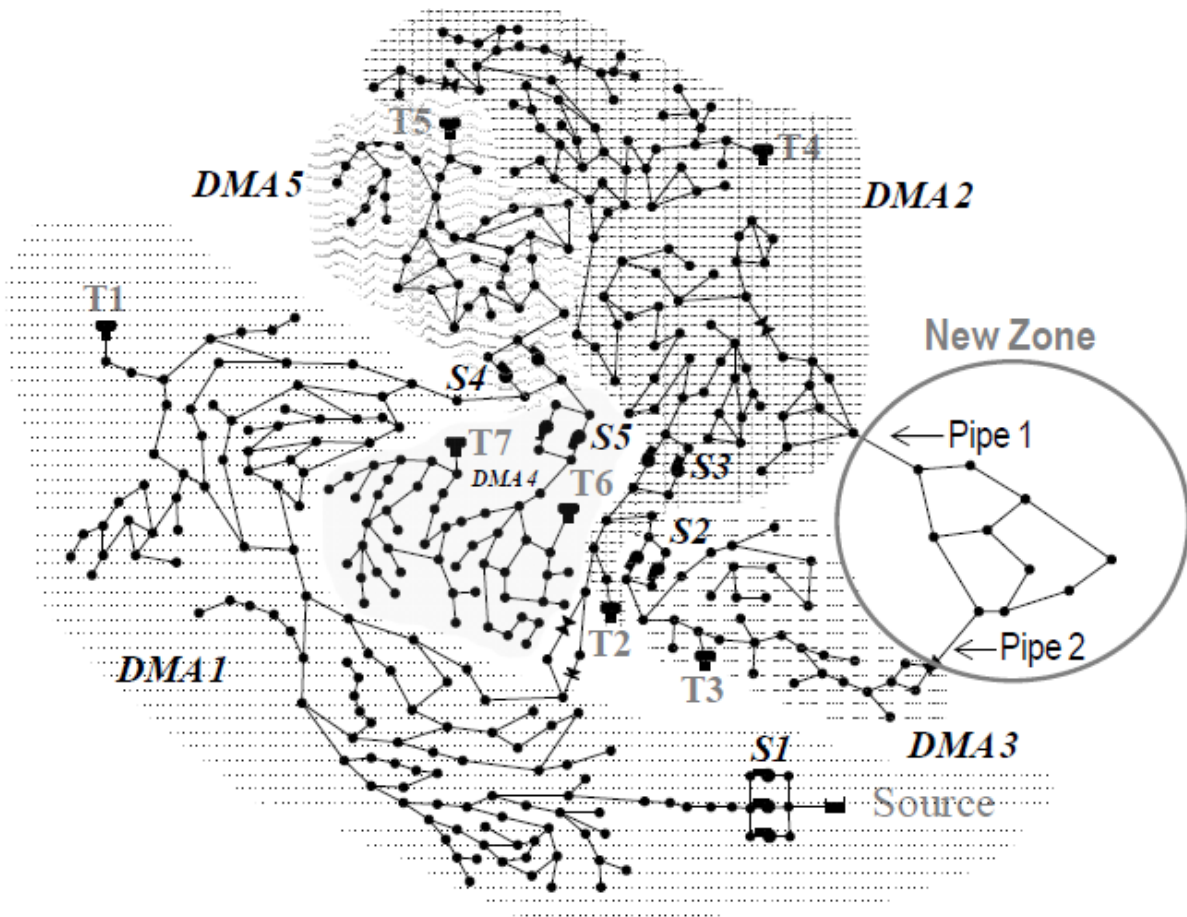


Figure 3.3 : D-Town network layout

The total cost requiring minimization is the sum of the annual capital costs and of the annual cost of pumping operations. The capital costs consist of component costs of pipes, pumps, valves, tanks and generators. The operational costs are calculated from the total system power usage under normal operating conditions based on a single design week. The electricity costs within the design week are specified according to normal peak and off-peak tariffs. The total GHG emissions include the emissions associated with the energy required for manufacturing, transportation and installation of the new pipes and the power usage from the operation of

pumps. The water age metric specified is WA_{net} as defined in Equation (3.8) with a water age threshold WA_{th} set to 48h.

Two operational scenario types for D-Town are specified, a normal operation scenario, for which the network was subject to normal demand loadings, and an emergency scenario, representing the event of a power failure. The design constraints for the normal operating scenario are distinguished as nodal constraints for the balancing period of a single design week. At each time point within this design week, the demand nodes are required to satisfy minimum head constraints, and the tanks are required to not empty. A hydraulic time-step of 15 minutes and a water quality time-step of 5 minutes are specified for the EPANET extended period simulations. The emergency scenarios are characterized by a power outage that can begin at any hour within the design week, and last for a duration of two hours. Within the emergency scenario, all pumps not powered by diesel generators are required to be shut down. Also constraints of minimum head for demand nodes, and non-emptying of tanks need to be met.

For the new zone, pipes are required to be sized from one of 12 diameter options (varying from 102 mm to 762 mm) for each link. The new zone is able to be connected via pipelines to either, or both, DMA2 and DMA3. For the pipe connection to DMA3, the design of a pressure-reducing valve (PRV) is permitted. The improvement options available to adapt the existing DMAs involved: addition of parallel pipes for all existing pipes (12 diameter options); increasing of storage volumes by one of six tank sizes (500 to 10,000 m³); addition of new pumps at the existing pumping stations (10 pump options are provided with varying head-discharge relationships); and sizing of backup power diesel generators for the pump stations (8 diesel generator options are available). For the existing DMAs, the valve settings for the existing valves are also allowed to be modified. In addition to the design options, operational pump scheduling decisions are also required to be made. As the network is specified to have a single week balancing period, the pump schedule for a single week needs to be determined. Operational controls are allowed to be either time-based, or based on threshold tank elevations.

3.2.4.2 Application results and discussions

3.2.4.2.1 Decision variables and bounds

Note that the order of the variables identified in this case study has no influence in solving the problem.

The first variable identified is the option of connection of the new zone to the existing network. This variable is modeled as an integer ranging from 1 to 5 for the five possible options which are: (1) DMA2; (2) DMA3; (3) DMA3 + PRV; (4) DMA2 + DMA3; and (5) DMA2 + DMA3 + PRV.

The application of Equation (3.1) to the new zone gives a diameter of 102 mm for a V_{max} of 3 m/s and a Q_{max} of 0.011 m³/s. Since 102 mm is the minimum allowed diameter option by the water utility, all the diameters are then set equal to 102 mm in the new zone. Therefore there is no variable associated with the pipe design in the new zone. This reduces the problem complexity by reducing the number of decision variables.

For the existing districts, the second set of variables identified is the pipe design option for each of the 429 pipes. Since cleaning and lining are not considered as options by the water utility, these variables are modeled as integers ranging from 1 to 3 for the three possible options which are: (1) do nothing; (2) duplicate; and (3) replace. The third set of variables is the diameter for each pipe in the existing districts. From the application of Equation (3.1) to the whole network, $D_{sup} = 406$ mm for a V_{max} of 3 m/s and a Q_{max} of 0.379 m³/s. These variables are modeled as integers ranging from 1 to $RD_{sup} = 7$ for the seven reduced possible diameter options which are (in mm): (1) 102; (2) 152; (3) 203; (4) 254; (5) 305; (6) 356; and (7) 406. Reducing the range of possible values for pipe diameter limits the search space and allows for a relatively rapid convergence. Note that if the chosen pipe design option is 1 (“do nothing”) in the optimization process, this variable is ignored in the decoding process.

The application of equations (2) to (4) gives three pressure zones for the D-Town network. Table 3.1 reports the computed characteristics for each pressure zone with $E_{min} = 3.48$ m, $E_{max} = 105.63$ m, $h_{min} = 25$ m and $h_{max} = 60$ m. However, in the D-Town network problem, the number of tanks, the tank bottom elevation and tank location are all fixed and the water utility does not

allow for changing these properties. The only change which can be applied to existing tanks is increasing their volumes (by adopting a bigger tank diameter) if needed.

Table 3.1 : D-Town Pressure zones characteristics

Zone	E_{min.zone} (m)	E_{max.zone} (m)	A_{min.zone} (m)	A_{max.zone} (m)
1	3.48	37.53	62.53	63.48
2	37.53	71.58	96.58	97.53
3	71.58	105.63	130.63	131.58

In order to evaluate the need for increasing tank volumes, equations (5) and (6) are engaged in different combination of districts. The results are reported in Figure 3.4 and Table 3.2 where NZ stands for new zone.

Table 3.2 : Reservoir volume and pumping capacity required for D-Town network

DMA considered	New Zone connection option	Q_{moy} (l/s)	Pumping station involved	Actual flow capacity (l/s)	Bs (m³)	Tank involved	Actual volume capacity (m³)
All	-	264	S1	79*3 = 237	2 906	T1 or All	5000 or 9500
2 + 3	4 or 5	93	S2 + S3	70 + 96 = 166	1 035	T2 or T3 + T4	2000 or 1500
3 – NZ	1	22	S2	35*2 = 70	237	T3	1000
3	2 or 3	30	S2	35*2 = 70	315	T3	1000
2	2 or 3	64	S3	48*2 = 96	729	T4	500
2 + NZ	1	71	S3	48*2 = 96	805	T4	500
5	-	32	S4	35*2 = 70	377	T5	500
4	-	38	S5	25*2 = 50	429	T6 + T7	500

DMA4 and DMA5 have been assessed separately because they were found to be independent from the rest of the network. Given different connection options of the new zone to the network,

the DMA2 and DMA3 can be linked in some way. Note that the new zone has the same demand pattern as DMA3. So in Figure 3.4 or Table 3.2, the DMA3 includes the new zone. Table 3.2 shows that the pumping station S1 is not able to provide the required flow and therefore one more pump is needed in parallel to the existing pumps in S1. All other pumping stations have a capacity greater than the required flow and don't need upgrading. The capacity of the tank T4 is not sufficient compared to the balancing reserve which leads to the decision of adding 1000 m^3 to this tank. All other tanks have a capacity greater than the required volume and don't need upgrading. Therefore there is no variable associated with the tank design. This reduces the problem complexity by reducing the number of decision variables.

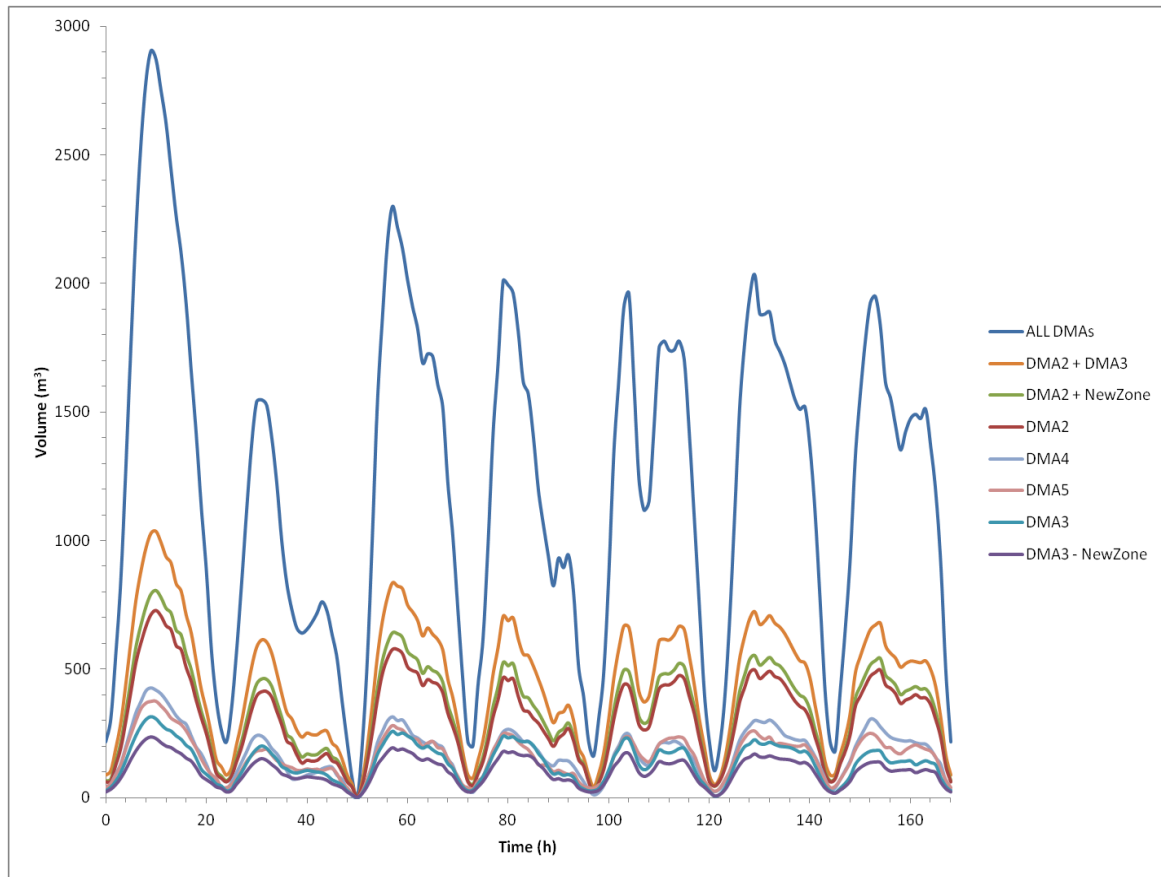


Figure 3.4 : Storage variation for different districts in D-Town network

The application results of Equation (3.7) to existing pumping stations are reported in Table 3.3. From these results, the pumping station S3 seems unable to provide the expected amount of

pressure to the whole demanding nodes of DMA2. But it should be noted, however, that the elevation of the tank T4 is sufficient to cover the required pressure. In light of these findings, all the existing pumping stations and tanks are found to be adequate enough in terms of producing required pressure. For pumping operation, a decision was made to use both time based and threshold tank levels based operation. The actual operation is exclusively based on threshold tank levels but it does not guarantee the constraint of the tank being at least half full at the end of the simulation week. This would lead in this instance to favoring a manual operation policy before the optimization process. The operation takes into account a proper filling and emptying for tanks and the constraint of tank's final level. Therefore there is no variable associated with pumps. This reduces the problem complexity by reducing the number of decision variables.

Table 3.3 : Pumping head required for D-Town

Pumping Station	DMA	E_{source} (m)	E_{max} (m)	$H_{\text{p.min}}$ (m)	$H_{\text{p.actual}}$ (m)
S1	1	59	46.16	12.16	40.56
S2	3	40	75.76	60.76	62.21
S3	2	40	105.63	90.63	86.60
S4	5	44	80.00	61.00	62.21
S5	4	44	79.05	60.05	64.18

The locations for valves are fixed in the network and only their setting can be modified. The actual settings of PRVs are 40m for each one and from the problem description the maximum setting should be 60 m. The fourth set of variables is then the valve setting which is modeled as a real ranging from 40 to 60 for each of the 4 PRVs. Depending on the connection option of the new zone chosen in the optimization process, the decoding of the variable associated with the PRV on the pipe 2 may be ignored.

From the analysis of Table 3.2, the pumping stations S2 and S4 can work very well with only one pump active. So diesel generators are provided for only pump in each of the two pumping stations. In the remaining stations, diesel generators are installed for all the pumps. In this way, the computation burden associated with power outage is reduced. Therefore there is no

variable associated with diesel generators. This is another means which reduces the problem complexity by reducing the number of decision variables.

Finally, the number of decision variables identified for the D-Town network optimization problem is 863: 1 for the new zone connection, 858 for existing pipes and 4 for pressure reducing valves.

3.2.4.2.2 Design constraints and objectives

The optimization objectives are clearly defined by the water utility for D-Town network: minimize total cost, minimize greenhouse gas emissions and minimize water age. Although three objectives are defined, the problem can be tackled as a single-objective optimization problem by combining all the objectives in one. However, this approach doesn't provide alternatives to enable compromise between different objectives. Therefore the multi-objective approach adopted in this article is meant to solve this problem. Given the decisions made in formulating variables, the total cost is the sum of two portions: a variable portion (pipe investment cost and pump operation cost) and a constant portion (valves, diesel generator, pumps and tanks investment cost). The constant portion is \$85,636. A penalty was added for (1) a negative pressure at non demanding nodes, (2) a pressure lower than 25 m at demanding nodes and (3) a tank final level lower than the initial level. Solutions with a zero penalty are awarded to those who fully respect the design constraints.

3.2.4.2.3 Optimization and results analysis

In order to determine the pay-off characteristic between the total greenhouse gas emissions and water age, the genetic algorithm was run with a population size of 30 sample solutions and was allowed to run for 500 generations. Approximately 9 seconds were spent to evaluate a single solution for the formulated D-Town network problem and therefore the whole optimization process expended approximately 37 computation hours on a virtual machine with a CPU 2.4GHz Intel Core i5. The determination to elect a relatively low value for the population size is due to the complexity of the D-Town network and the small time steps specified for hydraulic and water quality simulations. Figure 3.5 reports a portion of C++ code showing the way the genome was created for the D-Town network optimization problem. Note that the term “-Reliability” in

Equation (3.11) is replaced by “greenhouse gas emissions” for this problem. The power outage scenarios are only assessed at the end of the optimization process.

```

7055 // Create a phenotype for 863 variables.
7056 GABin2DecPhenotype map;
7057
7058 //New Zone connection option variable
7059 map.add(15, 1, 5.4);
7060
7061 //Pipe design options for the 429 existing pipes
7062 for (int i = 0;i<429;i++)
7063 {
7064     map.add(15, 1, 3.4);
7065 }
7066
7067 //Pipe diameters for the 429 existing pipes
7068 for (int i = 0;i<429;i++)
7069 {
7070     map.add(15, 1, 7.4);
7071 }
7072
7073 //Valve settings for the 4 PRVs
7074 for (int i = 0;i<4;i++)
7075 {
7076     map.add(31, 40, 60);
7077 }
7078
7079 // Create the template genome using the phenotype map.
7080 GABin2DecGenome genome(map);

```

Figure 3.5 : Genome definition for the optimization of D-Town network

The Pareto front identified by the GA at the end of iterations contained 11 solutions. The design values for seven of these solutions are summarized in Table 3.4 where V1, V45 and V47 are the three existing valves in DMA2 and N15 is the new valve on pipe 2. The seven solutions are pretty similar in terms of design options and objective function values. No multiple decision method is in need here to analyze this small number of solutions. The total cost is mainly influenced by the pipe design. The number of parallel pipes (duplicated or replaced) ranges from 239 to 246. From the analysis of the results, the solution (4) is recommended for the benefits of attaining the lowest cost, greenhouse gas emissions and number of parallel pipes. The only

concern ensued by this solution is amassing the highest water age metric. It is important to note, however, that all the water age values are closely ranked with each other.

Table 3.4 : Summary of design values for the seven selected solutions for D-Town

Solution number	1	3	2	4	5	6	7
Total Cost (\$/year)	626 819	632 458	635 634	624 806	634 744	639 738	632 744
GHG ($10^6 \text{kgCO}_2\text{-e/year}$)	3.32	3.33	3.33	3.31	3.36	3.39	3.34
WA _{net} (h)	0.21	0.21	0.20	0.21	0.16	0.15	0.20
New zone connection	5	5	5	5	5	5	5
No. duplicated pipes	170	172	177	171	180	178	170
No. replaced pipes	70	69	66	68	66	67	72
Setting PRV V1 (m)	56	56	56	56	56	55	56
Setting PRV V45 (m)	49	49	49	50	49	49	49
Setting PRV V47 (m)	55	55	55	55	55	55	55
Setting PRV N15 (m)	54	54	54	54	54	54	54

The reduction of the number of decision variables and their variation range was a helpful factor in tackling this complex problem. From the competition results obtained at the conference in Adelaide, the D-Town network design if formulated purely as an optimization problem, the span of the search space could easily reach over 7,500 decision variables depending on the options considered which implies that a large amount of time would be needed to solve the problem in such a situation. Through extensive preliminary analysis which resulted in uncovering four sets of decision variable groups, this greatly narrows the field to only 863 decision variables which highlights the potential of approaching complicated design issues by reducing the degree of complexity to solve the problem. It is obvious that this approach does not guarantee the global optimum as is the case for any approach intended to solve these kinds of problems. However, it does provide reasonable solutions in reasonable times to help in meeting this prominent challenge decision-makers face with real-world design problems.

Table 3.5 reports the objective function values found in the Battle compared to solution 4. As can be seen, the seven authors above the solution 4 in the mentioned table found more performing solutions than the one found in this paper. The rest of the authors found a less performing solution regarding the water age objective (the five authors below solution 4) or the total cost objective (the last two authors). However, the optimization problem was not tackled in this paper as in the context of the Battle. The aim was to test the applicability of an overall formulation approach on a specific design problem as complex as the one addressed in the conference.

Table 3.5 : Objective function values found in the Battle compared to solution 4 (data from Marchi *et al.* 2013)

Authors	Total Cost (\$/year)	Total GHG (kgCO ₂ -e/year)	WA _{net} (h)
Alvisi et al.	410 414	2 278 017	0
Guidolin et al.	420 410	1 588 458	0
Walski	424 446	2 276 659	0
Matos et al.	523 682	2 040 622	0.059
Wang et al.	385 777	2 237 599	0.095
Wu et al.	432 900	2 183 932	0.114
Tolson et al.	356 639	1 922 533	0.145
Solution 4	624 806	3 314 510	0.21
Saldarriaga et al.	433 790	2 003 077	0.229
Kandiah et al.	341 717	2 063 490	0.31
Iglesias-Rey et al.	378 860	2 055 239	0.612
Morley et al.	578 218	1 998 674	0.636
Bent et al.	396 723	2 539 008	1.099
Stokes et al.	922 421	2 733 235	0.127
Yoo et al.	928 227	2 172 386	0.193

The aim of the conference, as reported in the Battle paper (Marchi et al. 2013), was not to identify the best approach to solve WDS problems. Indeed, a comparison would not be possible given the diversification of resources used by the participants. Moreover, the run time was not a criterion considered in the competition. The aim of the conference was to identify approaches, challenges and results in solving the design problem in order to draw conclusions. A major conclusion from the conference was: there is benefit in using a combination of engineering experience and optimization methods when tackling complex real-world WDS design problem. However, the conference does not provide guiding on the aspects of approaches that can be generalized to other networks. The formulation approach proposed in this paper has the advantage of being applicable to multiple networks and it provides a competitive solution regarding the results of the conference.

3.2.5 Conclusion

A step-by-step approach for WDS design and operation has been presented, emphasizing on the need for a preliminary analysis before the optimization step, the reduction of the number and the range of decision variables, and the generalization of the design approach. In order to enlarge the scope of application of the proposed approach, some parametric equations were provided. The approach was applied to a large and complex network presented for a competition in the recent Water Distribution System Analysis (WDSA) conference in Adelaide, South Australia. The approach was able to formulate the problem with only 863 decision variables compared to the search space which could easily reach over 7,500 decision variables depending on the options considered. Therefore, the proposed approach has a potential in solving WDS design problems. However, given the diversity of constraints which could be required for each network, further applications on various networks may provide insights regarding ways to improve the proposed approach. The process could very well be applicable in future especially on WDSs with more flexible terms of design constraints. It can be made to test the other aspects of the proposed approach such as tank location or tank elevation design.

CHAPITRE 4 ARTICLE 3: MULTI-OBJECTIVE STOCHASTIC DESIGN OF WATER DISTRIBUTION SYSTEMS INCLUDING WATER QUALITY

4.1 Présentation de l'article

Dans cette section, l'objectif de l'article sera présenté, ainsi qu'une conclusion sur ce travail dans le cadre de la thèse. Cet article adopte une approche de formulation similaire à l'article 2 (voir chapitre 3) à la grande différence que le problème de conception multiobjective des réseaux de distribution d'eau potable (RDEP) incluant les réservoirs est traité ici dans un environnement stochastique. Le modèle proposé peut être utilisé pour la conception optimale d'un nouveau réseau ou l'extension d'un réseau existant, moyennant une intervention dans le code source du programme développé. L'approche se veut également un bon outil d'aide à la décision en proposant des solutions avec le risque de défaillance associé.

4.1.1 Objectif

Le caractère aléatoire du comportement des consommateurs ou des facteurs externes pouvant influencer la conception des RDEP est un aspect inévitable dans tout projet visant l'optimalité de cette conception. L'approche proposée dans cet article, contrairement à celui de l'article 2, est explicitement implémenté dans un environnement stochastique pour tenir compte directement des incertitudes liées aux données d'entrée. Les demandes aux nœuds sont alors traitées comme des variables aléatoires obéissant à une loi de probabilité donnée. Ces variables aléatoires sont supposées corrélées entre elles avec un certain coefficient de corrélation. L'algorithme d'optimisation est implémenté en deux étapes : une première étape déterministe suivie de l'étape stochastique pour accélérer la convergence. L'applicabilité de l'approche est testée sur le réseau de la ville de Saint Jérôme (65 000 habitants) pour évaluer son efficacité sur un réseau réel et complexe et dans lequel la qualité de l'eau joue un rôle important dans le processus d'optimisation.

4.1.2 Conclusion

La prise en compte explicite des incertitudes liées aux données est exigeante en termes de ressources de calculs. Elle exige alors énormément de temps, comparée à l'approche déterministe. L'implémentation de l'algorithme en deux étapes a permis de réduire ce temps de calcul. La présente approche a été capable d'identifier une solution plus fiable pour l'ensemble des fonctions objectives et plus particulièrement la qualité de l'eau comparée aux résultats des précédentes études sur le même réseau. Toutefois, bien que n'étant pas la plus réaliste, l'approche déterministe propose des solutions qui peuvent être obtenues en beaucoup moins de temps comparée à l'approche stochastique qui peut s'avérer non applicable sur des réseaux trop exigeants en termes de temps de simulation.

Cet article est soumis à "Journal of water resources planning and management".

Batchabani, E. and Fuamba, M. (2014a). "Multi-objective stochastic design of water distribution systems including water quality." Journal of water resources planning and management, Submitted.

4.2 Multi-objective stochastic design of water distribution systems including water quality

Essoyeke BATCHABANI, Ph.D. Candidate, Department of Civil, Geological and Mining Engineering, École Polytechnique de Montréal, P.O. Box 6079, Station CV, Montreal, Quebec, Canada, H3C 3A7. Email: essoyeke.batchabani@polymtl.ca

Musandji FUAMBA, P.Eng. Ph.D., Professor, Department of Civil, Geological and Mining Engineering, École Polytechnique de Montréal, P.O. Box 6079, Station CV, Montreal, Quebec, Canada, H3C 3A7. Email: musandji.fuamba@polymtl.ca

4.2.1 Abstract

This paper describes a procedure for the robust design of Water Distribution Systems (WDSs) which incorporates the uncertainty and the correlation issues of nodal water demands in a multi-objective optimization scheme aimed at minimizing costs, minimizing water age and maximizing a robustness parameter which is defined here as the probability of satisfying minimum pressure head constraints at all nodes in the network. The proposed model is an extension of previous research works on the stochastic design optimization of WDS to include water quality as objective in the optimization process. This is especially useful for WDS including tanks since this is usually the case for most real world networks. Decision variables involved in the optimization process are problem specific and may include those of the pipes, tanks and pumps in the network. The uncertainty in nodal water demands is considered by using “Latin Hypercube Sampling” (LHS) technique with the assumption of beta probability density function. The retrieved LHS samples are reorganized to include a certain correlation in nodal demands. The optimal design problem is solved using a standard Genetic Algorithm (GA) with a small number of samples for each fitness evaluation. The proposed approach is applied to a complex real-life case study dealing with the introduction of a tank in a medium-size city (65 000 inhabitants). The conclusive results’ analysis shows that the model produces more effective solutions in terms of reliability and water quality, with the use of an elevated tank operating with only one pump leading to significant cost savings compared to a ground tank which would require two pumps in this specific case.

Keywords: uncertainty, multi-objective, genetic algorithm, water distribution systems, optimization, water quality, Latin hypercube, correlation.

4.2.2 Introduction

Water Distribution Systems (WDS) play an important role in supplying sufficient quantity of potable water to consumers. These systems are designed to deliver water from a source in the required quantity, quality and at satisfactory pressure to individual consumers. To achieve these goals, WDS are equipped with multiple components such as pipes, pumps, valves, storage reservoirs and tanks. Although a given WDS may not include all these components, a real network usually does have them all included in different proportions depending on its size. The main challenge is how to make the determination in sizing different components of a WDS and how to operate it so as to meet unique objectives and constraints. WDS design and operation would then be tackled as an optimization problem by using a Single-Objective (SO) or a Multi-Objective (MO) approach. According to Filion (2008), many practitioners and researchers now believe that SO deterministic design should give way to a more comprehensive MO stochastic design of water networks since the latter affords the designer an opportunity to balance various goals and objectives that more closely suit the needs of a municipality. Only recently researchers have started to tackle the MO stochastic design problem to account for the probabilistic nature of demands and other loads in WDS.

The stochastic least cost WDS design problem was first conceived and solved as a SO formulation by Lansey *et al.* (1989b) by applying normal distribution on nodal demands. The novel aspect in their work laid with the inclusion of uncertainties in required demands, required pressure heads, and pipe roughness coefficients. The uncertain parameters were treated as being part of design constraints instead of design optimization objectives. An SO formulation was also used in the model by Duan *et al.* (1990) where the reliability associated with pumping stations was analyzed. Tolson *et al.* (2004) combined a Genetic Algorithm (GA) with FORM (First-Order Reliability Method) the WDS least cost design problem with water demands and roughness coefficients as uncertain inputs. In their model, they introduced correlation aspects in the uncertain inputs and concluded that correlations between random variables significantly increase optimal network costs. Babayan *et al.* (2005) optimized WDS using a Genetic

Algorithm (GA) to minimize total construction cost under the uncertain nodal demands sampled using Monte Carlo method from normal probability density function (PDF). In order to significantly reduce run times, Broad *et al.* (2006) implemented the use of Artificial Neural Network metamodels instead of EPANET to this design problem with uncertain water demands. More recently, Babayan *et al.* (2007) tested two different approaches to tackle the design issue: (1) the redundant design with safety margins added to uncertain parameters and (2) the integration method. They concluded that even slightly more expensive solutions procured using the redundant design method can be obtained with significantly less computational effort than those employing the integration method. Despite the great diversity of methods used to solve the stochastic least cost design problem, this approach has a limitation in that the solution sacrifices robustness and/or reliability of systems. Reliability and robustness refer to the network's ability to supply consumers adequately under normal and exceptional conditions (Savic *et al.* 2008) even in incidents when a part of the network failed. This is the reason for the recent expressed interest in MO stochastic design of WDS in its regard to maximize reliability or robustness as well as to minimize cost.

Kapelan *et al.* (2005) optimized the WDS using MO GA considering nodal demands and pipe roughness as uncertain parameters which are sampled using "Latin Hypercube" (LH) method from uniform and normal distribution. Two objective functions looked at are the minimization of total cost and maximization of robustness. Filion *et al.* (2007b) went beyond the hydraulic performance of the system and presented a stochastic design approach that quantifies the expected annual damages sustained by residential, commercial, and industrial users during low- and high-pressure hydraulic failures in a WDS. Giustolisi *et al.* (2009) used Beta distribution to evaluate uncertainties on nodal demands and pipe roughness with the same two objective functions as Kapelan *et al.* (2005). The finite boundary of Beta distribution makes it more proper to represent those two uncertain parameters. Jung *et al.* (2010) applied the approach of Giustolisi *et al.* (2009) to solve the problem. The innovation in their study was the introduction of a Disturbance Index (DI) to compare the case where only critical node is used to calculate the nodal pressure variability versus using all nodes in network.

In the majority of the above mentioned studies, evolutionary search techniques have been deployed to solve the MO stochastic design problem and nearly all of them deal with a

maximum of three objectives. Water Quality is not included in the optimization process and its modeling requires an extended period hydraulic simulation to be executed. Generally, this requires a shorter time step and therefore more computations than steady-state modeling (Broad et al. 2006). This limitation prevents the implementation of the above models to WDS including tanks since the introduction of a tank in a WDS is coupled with water quality issues. Indeed, all the networks tested, namely Anytown Network (Babayan et al. 2005; Duan et al. 1990; Filion et al. 2007b; Lansey et al. 1989b), New York Tunnels Network (Babayan et al. 2007; Broad et al. 2006; Jung et al. 2010; Kapelan et al. 2005), Apulian Network (Giustolisi et al. 2009) and Two-Reservoir Network (Tolson et al. 2004) are tankless. Given the growing recognition of the value tanks add to WDN, analysis of different means for incorporating tank considerations into WDN optimization models is needed (Batchabani & Fuamba, 2014b). This paper extends the MO stochastic design of WDS to include tanks and therefore water quality in the optimization process. Three objective functions are contemplated in solving the problem. The first one is the minimization of the total cost whose evaluation is deterministic. The other two are the result of stochastic calculation: minimization of water age and maximization of network robustness. The aspects emphasized in this paper are: (1) the inclusion of water quality as part of the objective functions in the stochastic optimization process, (2) the integration of the correlations in the stochastic nodal demands, (3) the use of a relatively complex real-life case study including tank design, and (4) the demand for a preliminary analysis aiming to reduce the number and the range of decision variables before the optimization process. The methodology uses a genetic algorithm to solve the formulated optimization problem. The algorithm coded in C++ is coupled with the well known simulation model EPANET (Rossman, 1999). The design options and criteria discussed in this article are not comprehensive but engineers frequently encounter them in real practice.

The next section of this paper describes the proposed methodology which a WDS designer could adopt to deal with the stochastic optimization problem. Results from the application of the approach on a complex real-life case study are presented in the section right after. Finally, the last section offers a conclusion and a recommendation for future applications.

4.2.3 Methodology

The methodology proposed to achieve WDS stochastic optimization follows different steps from the formulation to the resolution. This section begins by identifying and formulating the main variables associated with the WDS stochastic design, and the following subsection defines the performance criteria which may be used in the optimization process. The last subsection then introduces the method proposed in this paper to solve the formulated multi-objective stochastic optimization problem.

4.2.3.1 Identify and formulate the decision variables and set bounds

The objective of this section is to identify and formulate the main variables usually associated with the WDS stochastic design. Four network components are considered to be the main sources of the design variables. Each of these components with the associated variables is described in the following subsections.

4.2.3.1.1 Pipes

The first variable of each pipe considered could be the design option. Several options may be noted depending on the problem nature (new design, rehabilitation or network expansion). These options include: (1) cleaning, (2) lining, (3) duplicating, (4) replacing, or (5) doing nothing. The “option” variable can be modeled as an integer, ranging from 1 up to the number of the considered options. For a new design problem, this variable can be ignored.

The second variable for each pipe is the diameter size to be used. This diameter has to be chosen in a list of discrete commercial diameters. Although this list may be extensive, it is important to reduce the variable range in order to guide the optimization process toward a reasonable solution space. In this approach, a diameter is computed to be the upper limit for the diameter size. This diameter (D_{sup}) is based on a maximum velocity V_{max} imposed by the designer. The maximum water demand Q_{max} for a single time step Δt has to be computed based on the deterministic loading scenarios data. This water demand is increased by a factor β greater than one to account for the stochastic demand variation. The upper limit diameter (D_{sup}) is then given in Equation (4.1) where Q_{max} and V_{max} are expressed in SI. D_{sup} can be interpreted as the minimum diameter of a pipe that would carry Q_{max} at a maximum velocity of V_{max} .

$$D_{sup} (mm) = \text{Nearest Upper Commercial Diameter} \left(1000 * \sqrt{\frac{4 * \beta * Q_{max}}{\pi * V_{max}}} \right) \quad (4.1)$$

The commercial diameters have to be sorted in ascending order with a rank assigned to each size. If RD_{sup} denotes the rank assigned to D_{sup} , the “diameter size” variable can be modeled as an integer, ranging from 1 up to RD_{sup} . For a network rehabilitation problem where the pipe sizes cannot be modified, this variable can be ignored.

4.2.3.1.2 Tanks

There are several variables which could be theoretically associated with tanks in WDS design optimization (Batchabani & Fuamba, 2014b). In practice, these variables are highly case specific. With this approach, some parameters are computed in order to guide a decision regarding the usage of tanks.

If the number of tanks is unknown, the approach suggests computing the number (N_{zones}) of pressure zones. According to Walski (2000), pressure zone hydraulic grade lines should differ by roughly 30.48 m from one pressure zone to the next. Based on this assumption, N_{zones} is given by Equation (4.2) where E_{min} and E_{max} are respectively the minimum and maximum node elevations of the network. The suggestion made is to provide one tank in each pressure zone, giving a number of potential tanks equals to N_{zones} .

$$N_{zones} = \text{Round} \left(\frac{E_{max} - E_{min}}{30.48} \right) \quad (4.2)$$

If the tank bottom elevation variable has not been predetermined, the designer needs to set the bounds on this variable for the optimization process in order to reduce the search space. The recommended bounds are provided in each pressure zone in Equations (4.3) and (4.4) based on the elevation of the pressure zone that the tank will serve. In these equations, $A_{min,zone_i}$ and $A_{max,zone_i}$ are respectively the minimum and maximum tank bottom elevation in the pressure zone i ; $E_{min,zone_i}$ and $E_{max,zone_i}$ the minimum and maximum node elevations of the pressure zone i ; h_{min} with h_{max} the minimum and maximum allowable pressures. These bounds are computed so that the tank could give enough pressure to the highest customer and non-excessive pressure to the lowest one. The “tank bottom elevation” can be modeled as a real variable, ranging from $A_{min,zone_i}$ to $A_{max,zone_i}$ in each pressure zone.

$$A_{min.zone_i} = E_{max.zone_i} + h_{min} \quad (4.3)$$

$$A_{max.zone_i} = E_{min.zone_i} + h_{max} \quad (4.4)$$

Tank volume is relegated by the capacity of components such as reserves for production, emergency, fire or balancing depending on the use of the storage. A tank should basically supply the balancing reserve for useful storage. Emergency, production and fire reserves can be estimated based on local guidelines. This article offers an approach to estimate the balancing storage $B_s(m^3)$ before moving forward to the optimization step. The estimation for this reserve is based on a full time pumping during the entire extended-period simulation time. The approach has to be applied on the whole network and on the independent pressure zones in order to evaluate the balancing reserve for different tanks. The steps to follow are: (1) calculate the average water demand $Q_{moy}(m^3/period)$ which represents the uniform pumping rate along the extended period simulation; (2) compute the tank storage volume for each time period i , $S_i(m^3/period)$ by utilizing the continuity principle as in Equation (4.5) with an initial storage S_0 of zero; (3) the balancing storage is then given by Equation (4.6). Q_i in Equation (4.5) denotes the total water demand ($m^3/period$) in the period i for the considered pressure zone.

$$S_i = \beta(Q_i - Q_{moy}) + S_{i-1} \quad (4.5)$$

$$B_s = \max_i S_i + |\min_i S_i| \quad (4.6)$$

Evaluating the balancing reserve coupled with the uniform pumping rate is useful especially for defining if the existing pumps and tanks are sufficient or in need of upgrading. It is also important for the optimization process to reduce the number and the range of decision variables. With the estimated needed volume known, an assumption has to be made on the shape (diameter-height ratio and free volume ratio) to set the tank diameter. The diameter-height ratio is the ratio between the tank diameter and the tank total height, while the free volume ratio denotes the ratio between the free volume of the tank and its total volume. This free volume is the safety volume of the tank used only in the case of tank overflow.

The choice in determining tank location in real-world problems can be constrained by the availability of land already owned by the water utility or by the urban texture. In order to define this potential variable, specific information is required. If several nodes are regarded potential

points for a tank location in a given pressure zone, this variable for the corresponding tank can be set as an integer starting from 1 up to the set number of candidates.

4.2.3.1.3 Pumps

There are two types of variables for pumps: pump design variables and pump operation variables. For pump design, the two parameters involved are the flow and the head. The minimum required flow is βQ_{moy} (in Equation (4.5)) and the minimum required head $H_{p,min}$ (if head losses are not assumed) is given in Equation (4.7) where E_{source} signifies the elevation of the pumping source. The couple $(\beta Q_{moy}, H_{p,min})$ has to be regarded in selecting pump curves for each pressure zone.

$$H_{p,min} = E_{max} + h_{min} - E_{source} \quad (4.7)$$

For the available pump curves, it is important to compute the maximum power with the corresponding flow and head to ascertain if these curves meet the network needs or if it is necessary to combine them for this purpose. For a rehabilitation problem, these calculations help to figure out if actual pumps need to be upgraded or not.

Pump operation can be either time-based or based on threshold tank levels or both. If only time-based controls are chosen, the pump operation variables can be formulated as binary numbers corresponding to pump status (ON/OFF) in each time step. For the threshold tank levels based operation, the variables are the tank levels leading to pump switching (ON/OFF).

4.2.3.1.4 Valves

Different types of valves can be used to limit the pressure or flow at a specific point in the network. These types include pressure reducing, pressure sustaining, pressure breaker, flow control or general control valves. If valves are to be included, the decision variables could be the location and the setting. Typically, the location is imposed by the configuration of the network, leaving the setting as the only adjustable variable. For a rehabilitation problem, the existing setting may still be effective and no change would then be required.

4.2.3.2 Define and formulate the design performance evaluation criteria

In order to evaluate the design solutions characterized by assigning values to the variables discussed in the previous section, the performance assessment metrics have to be defined. Only feasible solutions are normally desirable and among these solutions, the best ones are found in the optimization process. This section begins with discussing the design constraints to define feasible solutions before dealing with design objectives which guide the optimization process.

4.2.3.2.1 Loading scenarios and design constraints

Loading scenarios with the corresponding design constraints, set by the designer, may include normal and abnormal operation conditions. For these scenarios to run effectively, the simulation must be balanced by a specified period that can be timed in an hourly, daily, weekly or any other balancing period. The hydraulic and water quality time-steps have to be selected based on the desired level of precision and the resulting computation burden.

The nodal demands are depicted in this approach as random input parameters for solving the problem. These demands are viewed to vary according some probability distribution function about nominal values which are the nodal baseline demands provided by the user. A study by McKay *et al.* (2000) comparing three methods for sampling input variables showed the robustness of the “Latin Hypercube Sampling” (LHS) compared to the other techniques in that it not only ensures that each of the input variables has all portions of its distribution represented by input values but also reduces the number of samples to be used for reliable statistical results. LHS is used in this study as a sampling method for nodal demands inputs.

Regarding the probability distribution function to be used, a recent study by Giustolisi *et al.* (2009) comparing the use of different distributions to reflect nodal demands variation revealed a preference for the Beta distribution especially for its finite boundary which makes it more proper to represent these uncertain parameters avoiding unrealistic extreme values. In addition, the Beta parameters can be chosen in a way that the generated values vary in a specific range defined by the user. Therefore, the Beta distribution does not need to be truncated to avoid for example generating negative values of nodal demands. The Beta probability density function (PDF) of variable x is shown in Equations (4.8) and (4.9).

$$f(x; a, b) = \frac{1}{B(a, b)} x^{a-1} (1-x)^{b-1} \quad (0 < x < 1) \quad (4.8)$$

Where

$$B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx \quad (4.9)$$

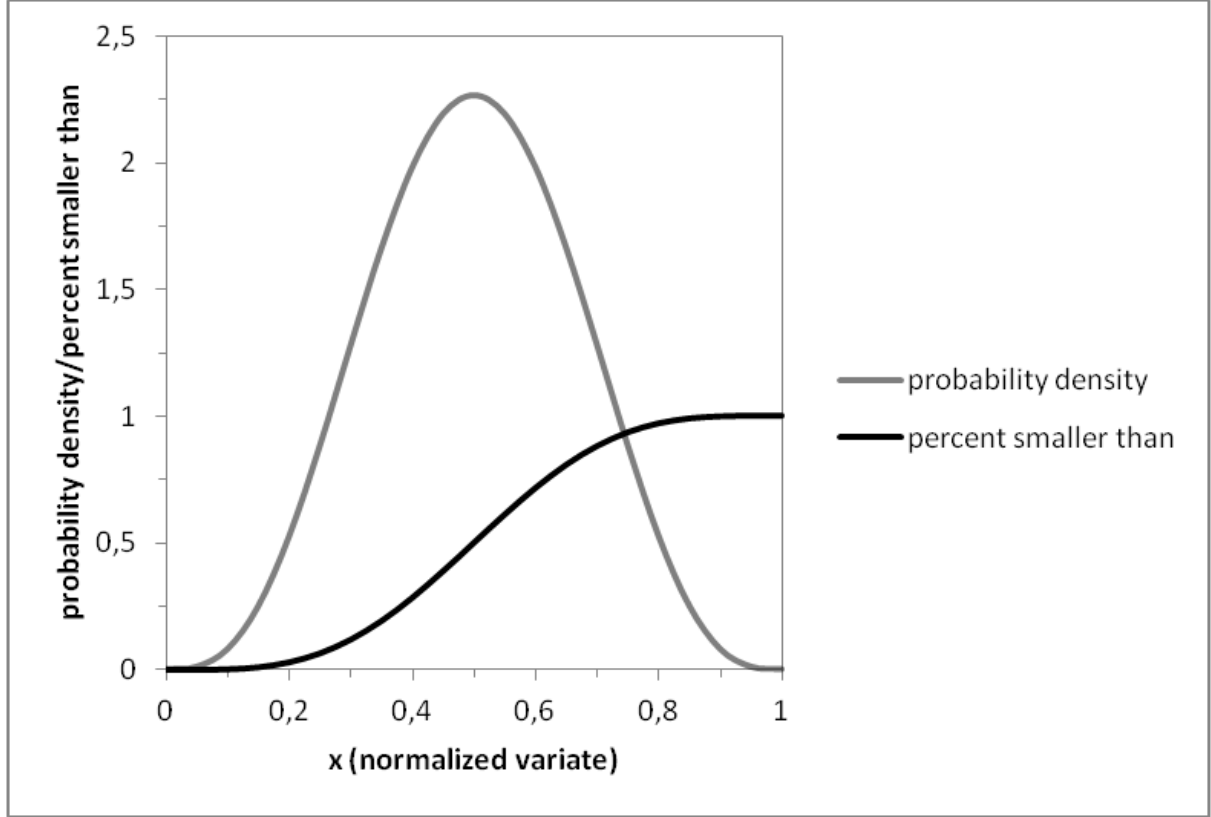


Figure 4.1 : Probability density and cumulative probability of Beta symmetric distribution

Beta symmetric is a Beta PDF ($a = b = 4.27$), having a bell shape similar to the normal PDF, assuming an average value equal to 0.5. Figure 4.1 shows the PDF and CDF (cumulative probability density function) of Beta symmetric distribution. This distribution is used in this study to consider uncertainty of nodal demands.

LHS is performed from the CDF. If N denotes the size of the samples to be used for the simulation, LH method samples a parameter from a region of CDF ($F(x; a, b)$) among the total N regions divided with equal probability. Sampled CDF value is converted to $(Q_{unc,i})$ using Equation (4.10).

$$Q_{unc,i} = Q_i + Q_i[F(x; a, b) - 0.5] \quad \in [Q_i - (Q_i * 0.5); Q_i + (Q_i * 0.5)] \quad (4.10)$$

Where Q_i is the baseline demand of node i , $i = 1, 2, \dots, N_{junc}$, the latter being the total number of demand nodes. From Equation (4.10), it can be clearly viewed that the coefficient of variation C_v assumed for baseline demands in this paper is 50%. The set of the N samples generated with this method doesn't include correlation of nodal demands. But in practice, some external factors may affect the system as a whole resulting in an increased or decreased demand at all nodes simultaneously. This is the motivation behind the integration of correlation issues in the samples. Filion *et al.* (2007a) also show how the cross correlation of demands can impact the water distribution network design. The samples generated above are then reorganized to include correlations in nodal demands. The procedure used for this is the one described by Kapelan *et al.* (2005) which was originally suggested by Iman and Conover (1982). If \mathbf{X} denotes the matrix of independent random samples generated by the LH method and \mathbf{R}^* the matrix of target correlations, the main steps of the Iman and Conover's procedure are as follows:

1. Calculate the actual correlation matrix \mathbf{R} using the data from \mathbf{X} . Each element in matrix \mathbf{R} represents the correlation coefficient calculated using the two data series from the corresponding rows and columns of matrix \mathbf{X} .
2. Decompose the target correlation matrix \mathbf{R}^* using Cholesky decomposition (Press et al. 1990). As a result, a lower triangular matrix \mathbf{P} is identified which satisfies the following equation $\mathbf{R}^* = \mathbf{P} * \mathbf{P}^T$.
3. Decompose the actual correlation matrix \mathbf{R} using Cholesky decomposition into $\mathbf{R} = \mathbf{Q} * \mathbf{Q}^T$.
4. Calculate the matrix $\mathbf{X}_1 = \mathbf{X} * (\mathbf{P} * \mathbf{Q}^{-1})^T$.
5. Create matrix \mathbf{X}^* by rearranging the columns in matrix \mathbf{X} so that the ranks in each column in \mathbf{X} are equal to the corresponding ranks in \mathbf{X}_1 . This way, matrices \mathbf{X}^* and \mathbf{X}_1 will have the same rank correlation matrix, and as a consequence, similar correlation matrices. \mathbf{X}^* contains the correlated nodal demands to be used in the stochastic simulations.

In this paper, only nodal baseline demands are assumed to be stochastic. The diurnal demand pattern (tendency of temporal variation) is assumed to be known. However, a demand with a stochastic tendency of temporal variation may be considered.

The design constraints can be measured in terms of minimum and maximum pressures provisioned at demand nodes for each time step within the design balancing period. Depending on the designer requirements, it is possible for many other constraints to develop. Handling constraints compels that a penalty function is used, so that infeasible solutions are penalized depending on their level of infeasibility, while feasible solutions come up with a penalty score of zero. Each constraint violation that appears is reported in this penalty function. However, in a stochastic environment, input loads are random and therefore output results are also random. This makes it difficult to distinguish feasible and infeasible solutions. The mean values are then used as reference for solutions evaluation.

4.2.3.2.2 Design objectives

Design objectives are the functions that can be maximized or minimized in the optimization process. WDSs are normally built and operated so as to: minimize capital and operation cost, maximize water quality, and maximize various system reliability metrics that include environmental considerations. In a stochastic environment, the reliability metric is referred in terms of a robustness parameter. The three targeted objectives in this paper are described in the next subsections.

Costs

The total cost includes capital costs and operational costs. Capital costs consist of component costs of pipes, tanks, pumps, valves and other possible problem specific components. The operational costs, on the other hand, are basically associated with the system power usage during the design horizon; however, other problem specific operation costs can be present. The costs must take into account the lifetime of each component and the discount rate. The pipe costs are function to the pipe diameter and length while tank costs basically subject to the tank volume. The pump investment costs are representative to pump power and the valve costs are inherent on their diameters. These costs are problem specific and part of input data. Since the nodal demands

are modeled as stochastic variables, the operational cost (which depends on demand) is stochastic. The cost objective could be defined in terms of minimizing the probability that the costs are greater than a specified target. This probability can be evaluated assuming a given probability distribution function as representative of the fluctuation of stochastic costs due to the demand uncertainty. In this paper, the cost objective is defined in terms of minimizing the expected value of total cost over the stochastic simulations.

Water Quality

Several parameters can be used to characterize the water quality: microbial growth, residual chlorine decay or hydraulic residence time. In numerical modeling, most of water quality models use the hydraulic residence time to compute different substance concentrations. This residence time which is expressed in the form of water age is used as water quality metric in the optimization process that can be defined in several ways. To account for water age impact on consumers, water age would be measured at non-zero demand nodes only while giving more weight to nodes with larger demands. The water age measure in Equation (4.11) includes these computations and can be then adopted in the optimization process.

$$WA_{net} = \frac{\sum_{i=1}^{N_{junc}} \sum_{j=1}^{N_{time}} k_{ij} * Q_{dem,ij} * WA_{ij}}{\sum_{i=1}^{N_{junc}} \sum_{j=1}^{N_{time}} Q_{dem,ij}} \quad (4.11)$$

In Equation (4.11), WA_{net} refers to the network water age, WA_{ij} tells the water age at junction i at simulation time t_j , k_{ij} is a binary variable defined as 1 if WA_{ij} is greater than a threshold WA_{th} and zero otherwise, $Q_{dem,ij}$ points to the demand at junction i at time t_j , N_{junc} is the number of system junctions and N_{time} counts the number of simulation time steps. Given the stochastic nature of the problem, there is a WA_{net} value associated with each single random loading sample. The stochastic value of water age used in the optimization process in this paper is the expected value (average) \overline{WA}_{net} across all the samples considered. However, the water age objective could be defined in terms of minimizing the probability that the water age is greater than a specified target. The statistic parameters of the stochastic water ages such as the mean and the standard deviation (dispersion) combined with a probability distribution function would provide the risk associated with water age for each set of design decisions.

Network robustness parameter

Assuming nodal demands as uncertain variables, according to a fixed PDF, means that the computed hydraulic output variables become stochastic. In order to maximize the probability of non-failure of the system, the paper defines a robustness parameter to be maximized. The robustness is defined here as the probability that nodal pressures are greater than the minimum requirements. More specifically, Equation (4.12) is considered in the optimization process.

$$\text{maximize Prob}(h_i \geq h_{\min}^0) \quad i = 1, \dots, N_{junc} \quad (4.12)$$

By taking into account the standard deviation and average value of stochastic outputs, assuming a normal PDF as representative of the fluctuation of stochastic outputs due to demand uncertainty, the following robustness parameter is considered (Equation (4.13)).

$$\text{maximize } \alpha = \min_i \alpha_i \quad \text{where} \quad \alpha_i = \frac{\bar{h}_i - h_{\min}^0}{\sigma_{h_i}} \quad i = 1, \dots, N_{junc} \quad (4.13)$$

As in Giustolisi *et al.*(2009), the pareto front is allowed to evolve between two extreme values of network robustness. The lower bound is 10% which implies that the relevant value of α is -1.282. The upper bound is set to 90% corresponding to α equal 1.282. The latter is imposed in order to confine robustness evaluation to what is technically suitable. Therefore, networks with robustness beyond this range are penalized during the optimization process.

4.2.3.3 Solve the multi-objective stochastic optimization problem

The objective of this section is to offer the method proposed in this paper to solve the multi-objective stochastic optimization problem formulated in the previous sections. The optimization and simulation tools are introduced and then the optimization process is presented. Finally, the optimization results are to be analyzed so as to select a proper design.

4.2.3.3.1 Genetic Algorithm

Although many advanced Genetic Algorithms (GA) do exist in the literature and problem specific genetic operators could be developed, the purpose of this work is not to promote a given GA. A steady state GA is used in this work and adapted for multi-objective optimization. For

this purpose, a GA library called GALib (Wall, 1996) is used, which provides necessary tools for GA and its customization. Also operators built in this library are used.

In order to accelerate the computation speed, Giustolisi *et al.* (2009) suggested the stochastic optimization is implemented in two stages. First, a deterministic problem is optimized with a large number of generations and then, the stochastic problem is optimized using the final solution from the first stage as a port of initial population. The latter is run for a small number of generations.

For the genome representation, the program uses a binary-to-decimal genome to handle the optimization variables. This genome is an implementation of the traditional method for converting binary strings to decimal values. To use this genome, what is required is the ability to create a mapping of bits to decimal values by specifying how many bits will be used to represent what bounded numbers.

4.2.3.3.2 EPANET solver

EPANET is executed here in command mode. A series of DOS commands are introduced in the code to run the simulation and save results into a text file. The saved report is a comprehensive report and the results recorded at each time step for the different variables are: (1) demand nodes: actual demand, total head, pressure, quality; (2) reservoirs: net inflow, total head, pressure, quality; (3) tanks: net inflow, total head, pressure, quality; (4) pipes: flow, velocity, unit headloss, status; (5) pumps: flow, quality, head, status (6) valves: flow, velocity, headloss, status. EPANET records daily energy consumption as well.

4.2.3.3.3 Pareto ranking

The score associated to the solution here corresponds to its rank in a given population (P) of solutions. This rank reflects the degree of dominance of the solution in this population. The constrained optimization problem is reduced to unconstrained optimization problem reported in Equation (4.14). Note the equivalence between maximizing α and minimizing $-\alpha$.

$$\min F(x) = (f_1(x) = Cost + Penalty, f_2(x) = \overline{WA}_{net}, f_3(x) = -\alpha) \quad (4.14)$$

In Equation (4.14), x denotes the set of problem variables discussed earlier. Each solution to the problem is associated with an objective vector $F = (f_1, f_2, f_3)$ to be minimized. Given two solutions to the problem x^1 and x^2 , x^1 is said to dominate x^2 if and only if $F(x^1)$ is partially less than $F(x^2)$; i.e., if and only if the conditions of Equation (4.15) are satisfied.

$$f_i(x^1) \leq f_i(x^2) \forall i = 1, 2, 3 \text{ and } \exists i \in \{1, 2, 3\} / f_i(x^1) < f_i(x^2) \quad (4.15)$$

A solution x^0 is said to be Pareto optimal or efficient with regard to a set of solutions if and only if there does not exist any solution in this set dominating x^0 . The set of non dominated solutions forms the Pareto front. The latter is the one analyzed by the decision maker to choose the solution it deems appropriate.

There are various Pareto ranking methods in a given population (Belegundu et al. 1994; Fonseca & Fleming, 1993; Goldberg, 1989). A comparative study by Mallor et al. (2003) showed that Fonseca & Fleming's method is stable and produces small errors in classification. This classification is the one used here to assign a rank to a solution x in P as shown in Equation (4.16). In this classification, the rank n of a solution x is the number of solutions that dominate x in P . The non-dominated solutions are, therefore, all of the same rank zero, while those dominated have ranks between 1 and $k - 1$, where k is the size of P .

$$Rank_{Pareto}(x) = \text{Number of solutions dominating } x \text{ in } P \quad (4.16)$$

For this purpose, a basic dominance function was created in the program. This function takes as arguments two vectors of three-dimensional real x^1 and x^2 . It returns a Boolean variable that is true when x^1 dominates x^2 according to Equation (4.15). Initializing the rank of x to 0, for each solution x^i in P , $i = 1$ to k , if the function dominance (x^i, x) returns true, the rank of x is incremented by one unit.

The score of each solution is set equal to its rank in the population. This score is to be minimized by GA. In this way, the GA will favor non-dominated solutions in each generation, allowing for fast convergence towards the optimal solutions. A population evaluator is then created in the program for this purpose. This function evaluates and sets the score of each genome in the current population during the GA evolving process.

4.2.3.3.4 Optimization process

The proposed program runs as follows (Figure 4.2):

- a. The program reads the input file: this allows recording relevant information on the network such as the number of demand nodes or nodal baseline demands.
- b. From (a), the program computes indicative parameters: pressure zones, pipe upper bound diameter, minimum and maximum tanks bottom elevations. This information is given to the user to assist him in setting decision variables and bounds.
- c. The program provides the GA the form of a genome using the decision variables set by the user from (b).
- d. The GA generates a set of random genomes in the same format that has been provided in (c). The number of genomes equals to the size of the population as computed earlier.
- e. The program retrieves each genome, converts its contents and saves it in the input variables. The LH provides the random nodal demand samples and for each resulting correlated sample, taken together with the converted genome information, the program writes EPANET's input file for simulation by EPANET solver.
- f. The program uses the stochastic simulation results from (e) to calculate the values of the objective functions and the penalty associated with mean respect to constraints.
- g. The program computes the Pareto rank of each genome from (f). These ranks are retrieved by GA. The ranking measures the quality of each genome.
- h. The GA uses genome scores from (g) to generate a new improved population.
- i. Steps (e) to (h) are repeated until the termination criterion is met.

The program saves the non dominated genomes into a text file. Values of objective functions and penalty are recorded as well in the same file. These results are then analyzed by the decision maker.

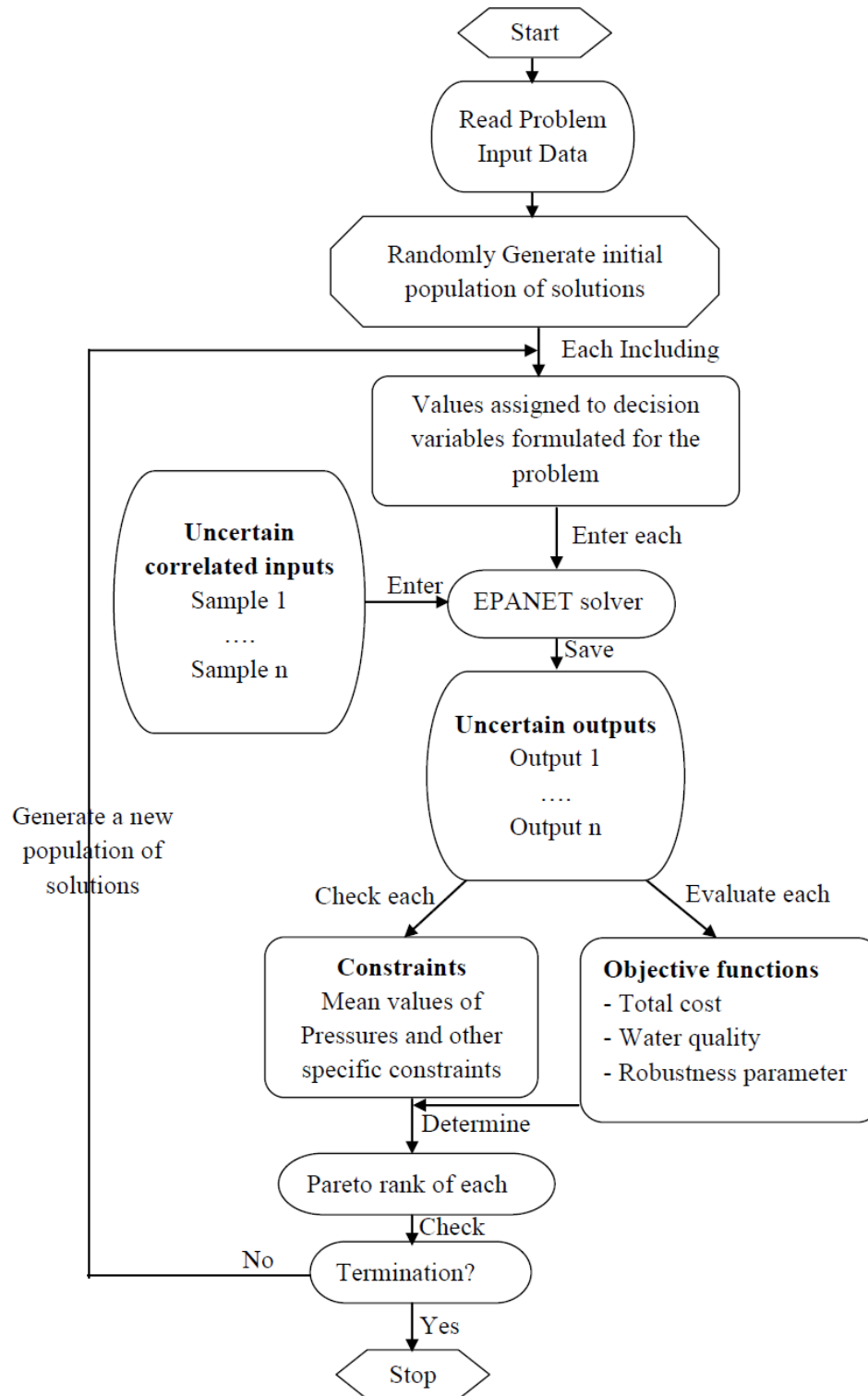


Figure 4.2 : Optimization procedure in the proposed methodology

4.2.3.3.5 Results analysis

The choice of one solution among the candidates from Pareto front is not universal and obvious. It depends on the decision maker's priorities since gaining on a goal deteriorates at least one other goal and vice versa. It then requires compromise from the designer.

The decision maker (DM) may set a maximum water age and then would reject solutions with \overline{WA}_{net} greater than this. If DM is less concerned about water quality and more demanding on robustness or total cost, his choice becomes more and more limited.

Pareto front is usually composed of a small number of solutions and the choice is relatively less polemical. If the difficult choice comes with a great number of non dominated solutions, one of multiple decision making methods such as AHP (Analytic hierarchy process) can be used. Pareto approach for multi-objective optimization allows for making a reasonable choice in a small space of solutions even if the choice is not necessarily the best.

4.2.4 Case study

4.2.4.1 Network description

The feasibility of the proposed approach is tested on the city of Saint-Jerome (Quebec, Canada) which is a typical medium-size municipality of 65 000 inhabitants, exhibiting a significant growth of around 1.4 % per year. The WDS of this city (Figure 4.3), a publicly owned and operated entity, is fed by one water treatment plant (WTP) located in a high pressure zone. Most of the treated water is gravity fed from the WTP to the lower pressure zones, exception made of a few areas (representing about 10% of the city), which are re-pressurized on the WDS. The node elevations range from 216.2 ft up to 639.8 ft, leading to several pressure zones in the system. The network has about 1500 nodes, 2000 pipes (280 km), 7 pumps and 48 control valves.

A forecast study carried out in 2001 suggests a lack of capacity to supply the future demand under several configurations (Pagé, 2005): (1) during the summer, the forecasted peak demand will be larger than the WTP capacity, and (2) the WDS will not be able to meet fire flow conditions for the south-east portion of the city. A fire flow of 5764 USGPM for a period of 5 hours was identified as the critical requirement for the south-east suburb. This flow is split

amongst 5 nodes around the building that was identified as the critical design location (Figure 4.3).

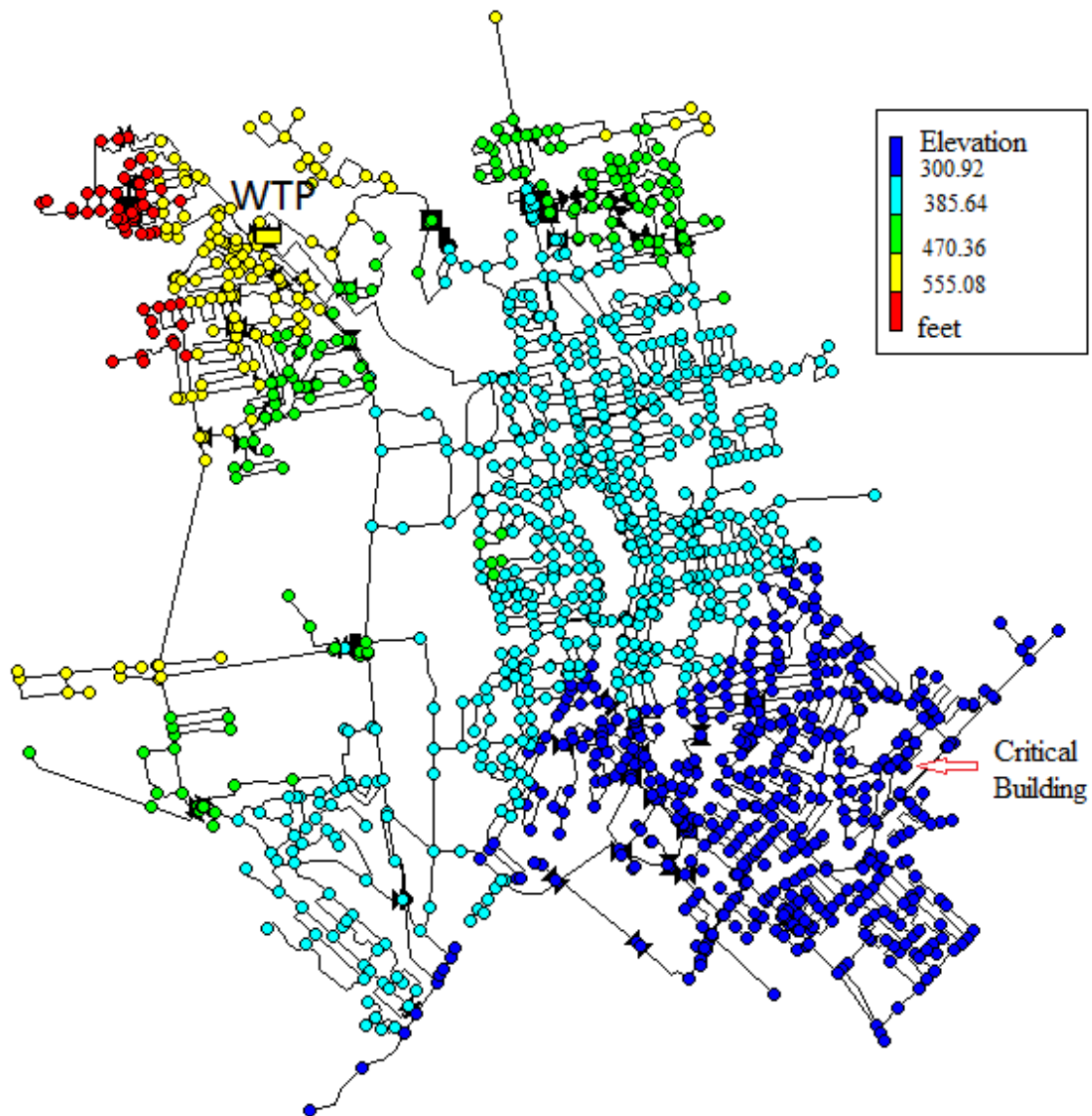


Figure 4.3 : Saint-Jerome WDS layout and elevations

Following those considerations, the WDS manager decided to install two new tanks: (1) A new tank located at the WTP built in 2008. This storage, with a capacity of 635 000 US gallons, increases the WTP equalization volume to 10 million US gallons. The new installation provides

sufficient storage for the forecasted peak demand. (2) A new tank to be constructed is designed to provide the fire flow for the buildings located in the south-east bound area.

The municipality of Saint-Jerome is in need of a tank design project to cope with these potential fire events. The water utility is naturally interested in minimizing operational and capital costs, but is also concerned about water quality issues in the network. The capital costs consist of the cost for the construction of the new tank and the investment costs of the eventual new pumps. The operational costs estimated are the result of the power usage of new pumps if applicable.

For Saint-Jerome, two operational scenario types have been identified, a normal operation scenario of which the network was subject to normal demand stochastic loadings, and an emergency scenario, representing the possible event of a critical fire. During the normal demand loadings, every demand node on the network is expected to have water delivered to it with adequate pressure. The minimum required pressure for nodes with demands is 20 psi while nodes without demand only have the requirement of a minimum pressure above zero. A typical day was broken into 24 one-hour sub-periods with the proportion of the average demand following the pattern of Figure 4.4. During the fire event, the system is still expected to provide normal demands at minimum required pressures and the tank would not be allowed to stay empty.

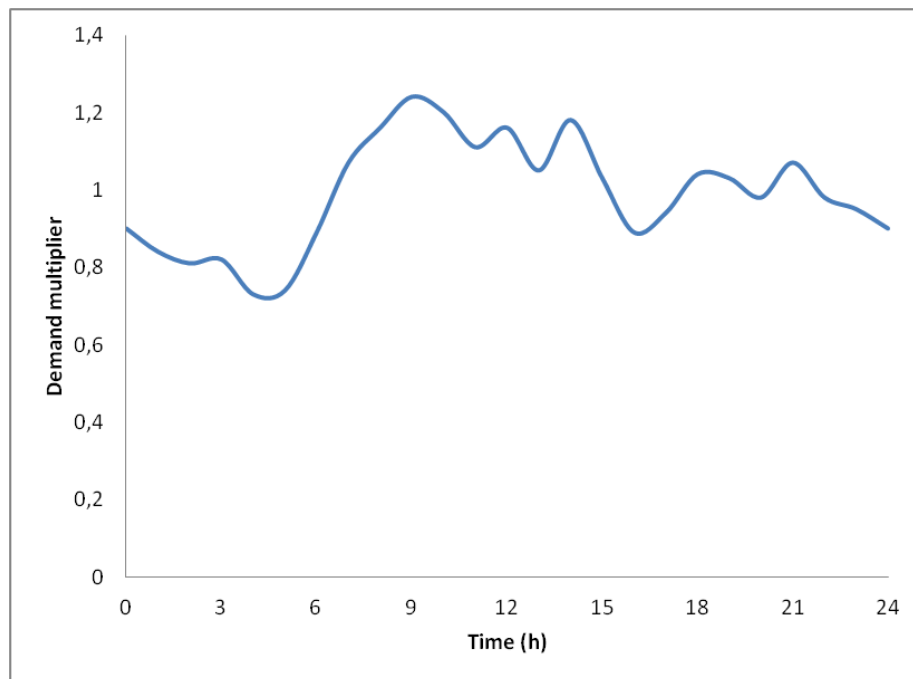


Figure 4.4 : Saint-Jerome network diurnal demand pattern

The type of the new tank, its location or shape are not specified and maybe therefore subject to optimization. This new tank is allowed to be coupled with new pumps and in this case the associated pumping operation is to be determined. Furthermore, the water utility does not allow for the existing components (pipes, pumps, valves) to be resized or for the current operation to be changed.

4.2.4.2 Application results and discussions

4.2.4.2.1 Decision variables and bounds

For this case study, there is no decision variable in relation to pipes or valves. The decision variables rather concern the tank and the potential pumps. The application of Equations (4.2) to (4.4) gives five pressure zones for the Saint-Jerome network. Table 4.1 reports the computed characteristics for each pressure zone with $E_{min} = 216.2 \text{ ft}$, $E_{max} = 639.8 \text{ ft}$, $h_{min} = 50 \text{ ft}$ and $h_{max} = 230 \text{ ft}$. In Figure 4.3, each pressure zone of the Saint-Jerome network is identified by a different color.

Table 4.1 : Saint-Jerome Pressure zones characteristics

Zone	$E_{min,zone} \text{ (m)}$	$E_{max,zone} \text{ (m)}$	$A_{min,zone} \text{ (m)}$	$A_{max,zone} \text{ (m)}$
1	216.20	300.92	350.92	446.20
2	300.92	385.64	435.64	530.92
3	385.64	470.36	520.36	615.64
4	470.36	555.08	605.08	700.36
5	555.08	639.80	689.8	755.08

The tank bottom elevation bounds reported in Table 4.1 are related to the elevated tank option if adopted. In order to be effective, the storage tank, designed to face a potential fire in the south-east portion of the city, should be within pressure zone 1. This pressure zone with the lowest elevations is identified by the blue color in Figure 4.3. It is assumed that any node in this pressure zone is able to get a new tank connection. So as to determine the type of tank to consider in this study, preliminary simulations had to be conducted. What it revealed was that an elevated tank without a pump would not be able to simultaneously fill and empty under normal

flow conditions given the hydraulic head of the WTP reservoirs. For low tank bottom elevations, the tank won't empty and for high elevations, it won't fill. One solution to this is to place the tank at a high elevation and use a pump to fill it, with the emptying simply carried out by the use of gravity. If a ground or buried tank option is selected, instead of an elevated tank, there would be a need for a pump to empty the tank. However, given the high fire flow, two different pumps would be essential in this case: the first one would be used to empty the equalization volume and the second one being significantly more powerful for the fire flow. With these considerations in mind, an elevated tank with one pump option is chosen for this study. The next paragraphs describe the selected variables and bounds. Note that the order of the variables identified in this case study has no influence in solving the problem.

The first decision variable is the tank location which is modeled as an integer ranging from 1 up to 524 (total number of nodes in the south-east pressure zone).

The second decision variable is the tank bottom elevation modeled as a real variable ranging from 350.92 ft up to 446.20 ft as reported in Table 4.1.

The diameter-height ratio of the tank is assumed in this paper to be 0.9 including a free volume ration of 10%. This gives a maximum operational height of the tank equal to its diameter which is determined by the tank operational volume (emergency + balancing). The third decision variable is the tank emergency volume (fire storage). The minimum fire volume is 120 m^3 as specified in the Quebec guidelines (MDDEP, 2006). This value is used to set the lower bound on this decision variable. The upper bound is computed using the deterministic fire flow (5764 USGPM) and duration (5 hours), increased by 50 % to account for the stochastic nature of the problem. The corresponding volume is leveled at about $9,800 \text{ m}^3$. The tank emergency volume is then modeled as a real variable ranging from 120 m^3 up to $9,800 \text{ m}^3$.

The fourth decision variable is the tank balancing volume. In order to limit this variable range, the equations (5) and (6) were applied with the factor β set to 1.5. The computed balancing volume measures to about $1,300 \text{ m}^3$. The upper bound is set to be the maximum volume needed for fire which is $9,800 \text{ m}^3$. The tank balancing volume is then modeled as a real variable ranging from $1,300 \text{ m}^3$ up to $9,800 \text{ m}^3$. These last two variables combined with the tank shape give the tank diameter and minimum/maximum operational levels.

The fifth and last decision variable is the pump flow. This flow serves only to fill the tank. The value of this variable is important given the fact that the duration required to fill the tank may affect the water quality. The pump flow is modeled as a real variable ranging from 50 USGPM up to 1000 USGPM which corresponds to the current pump flow range used in the Saint-Jerome network. The pump hydraulic head is set to be the difference between the tank elevation and the connection node elevation. Therefore, this hydraulic head is derived from the tank location and bottom elevation. For the new pump operation schedule, it was decided to base it on the new tank's threshold levels: the pump starts when the water level goes under the minimum level and stops when it goes up to the maximum level.

Finally, the number of decision variables identified for the Saint-Jerome network optimization problem is 5 where 4 relate to the tank design (location, bottom elevation, emergency volume, and balancing volume) and the remaining one is dedicated to the pump flow.

4.2.4.2.2 Design constraints and objectives

The optimization objectives are: minimize total cost (investment and operation), minimize water age (Equation (4.11) with a water age threshold WA_{th} set to 48h) and maximize the robustness parameter defined in Equation (4.13). Costs are calculated using the following hypothesis: an annual interest rate of 5%, an amortization period of 20 years (a typical amortization period in the construction field), a capital cost of \$3.78/USGal for the storage tank, a fixed cost for the pumping station of \$200,000, a monthly electrical cost of \$13.71 per kW of peak power (if higher than 50 kW) plus a tariff of \$0.0449 per kWh. These hypotheses are in line with current (2013) costs for the Hydro-Québec industrial electricity rates (Hydro-Quebec, 2013). The operational costs are updated on a yearly basis. The water age is computed during the normal demand scenario for a 240-hour stochastic simulation to reach stable solutions. The robustness parameter is computed during the critical fire scenario. The fire flow, split amongst 5 nodes around the critical building (Figure 4.3 and Figure 4.6), is treated as a random variable generated with the LH method using the beta distribution. During the fire scenario, the tank begins at its minimum operational level and the pump is off. A penalty was added for (1) a negative pressure at non demanding nodes, (2) a pressure lower than 20 psi at demanding nodes and (3) a tank running empty for a given time step during the fire scenario. Solutions with a zero penalty are awarded to those who fully respect the design constraints.

4.2.4.2.3 Optimization and results analysis

In order to determine the pay-off characteristic between the total cost, the robustness parameter and water age, the genetic algorithm was run with a population size of 30 sample solutions, and allowed to run for 100 generations in the deterministic stage with 10 generations in the stochastic stage. For the stochastic simulations, the used number of samples is equal to 30. This choice derives from the assumption that mean values and standard deviations evaluated on, at least, 30 samples can be reasonably considered as sufficiently statistically reliable (Benjamin & Cornell, 1970). The use of LHS method avoids using excessive high sample numbers which would result in excessively long computational times. During the process of generating correlated demands from LHS samples, the correlation coefficient between any two nodal demands is assumed equal to 0.50, as was done by Tolson et al. (2004) and Kapelan *et al.* (2005). Figure 4.5 reports a portion of C++ code showing the way the genome was created for the Saint-Jerome network optimization problem.

```

7003 // Create a phenotype for 5 variables.
7004 GABin2DecPhenotype map;
7005
7006 //Tank location variable
7007 map.add(15, 1, 524);
7008
7009 //Tank bottom elevation variable
7010 map.add(31, 350.92, 446.20);
7011
7012 //Tank emergency volume variable
7013 map.add(31, 120, 9800);
7014
7015 //Tank balancing volume variable
7016 map.add(31, 1300, 9800);
7017
7018 //Pump flow variable
7019 map.add(31, 50, 1000);
7020
7021 // Create the template genome using the phenotype map.
7022 GABin2DecGenome genome(map);

```

Figure 4.5 : Genome definition for the optimization of Saint-Jerome network

Approximately 20 seconds were spent to evaluate a single solution for the formulated Saint-Jerome network problem and the whole optimization process expended approximately 62 computation hours on a virtual machine with a CPU 2.4GHz Intel Core i5. The Pareto front identified by the GA at the end of iterations contained 3 solutions, all leading to the same design configuration reported in Table 4.2. This design consists of connecting the tank to the node 1012 (see Figure 4.6) at an elevation of 439 ft and would cost the Saint-Jerome water utility

\$8,566,000 dollars for the whole period of 20 years in operation. Note that 97% of this cost is attributed to the tank design.

Table 4.2 : Summary of design values for Saint-Jerome network optimization problem

Decision variables and objective functions	Value	Other resulting parameters	Value
Tank location	Node 1012	-	-
Tank bottom elevation (ft)	439	Pump head (ft)	200
Tank emergency volume (m ³)	4,300	Tank minimum operation level (ft)	40
Tank balancing volume (m ³)	3,200	Tank maximum operation level (ft)	70
Pump flow (gpm)	600	-	-
Total cost (10 ⁶ \$)	8.566	-	-
Water age (WAnet) (h)	0,99	Water age in the tank (h)	78
Robustness parameter (- α)	-0.84	Robustness (%)	80

Basile (2008) tried to solve the same design problem during his master thesis by studying only two tank location possibilities. The best solution found by his approach was to provide a ground tank connected to the node 902 (see Figure 4.6) and coupled with two pumps. His emergency volume was gauged at 6,000 m³ with an equalization volume of 5,300 m³. That yielded to a high water age of 110 h in the tank, which may lead to undesirable water quality degradation. The comparison of the solution identified in this paper and the one by Basile (2008) shows a certain number of good points regarding the current solution: (1) the tank volume is 65 % smaller; (2) the water age in the tank is 70 % lower; (3) the tank is located very close to the critical building where the fire was assigned (Figure 4.6); (4) the stochastic simulation of this problem leads to a reliable design with a robustness of 80 %, meaning that the solution has a chance of 80 % to satisfy the minimum pressure requirement at all nodes under random reasonable demands.

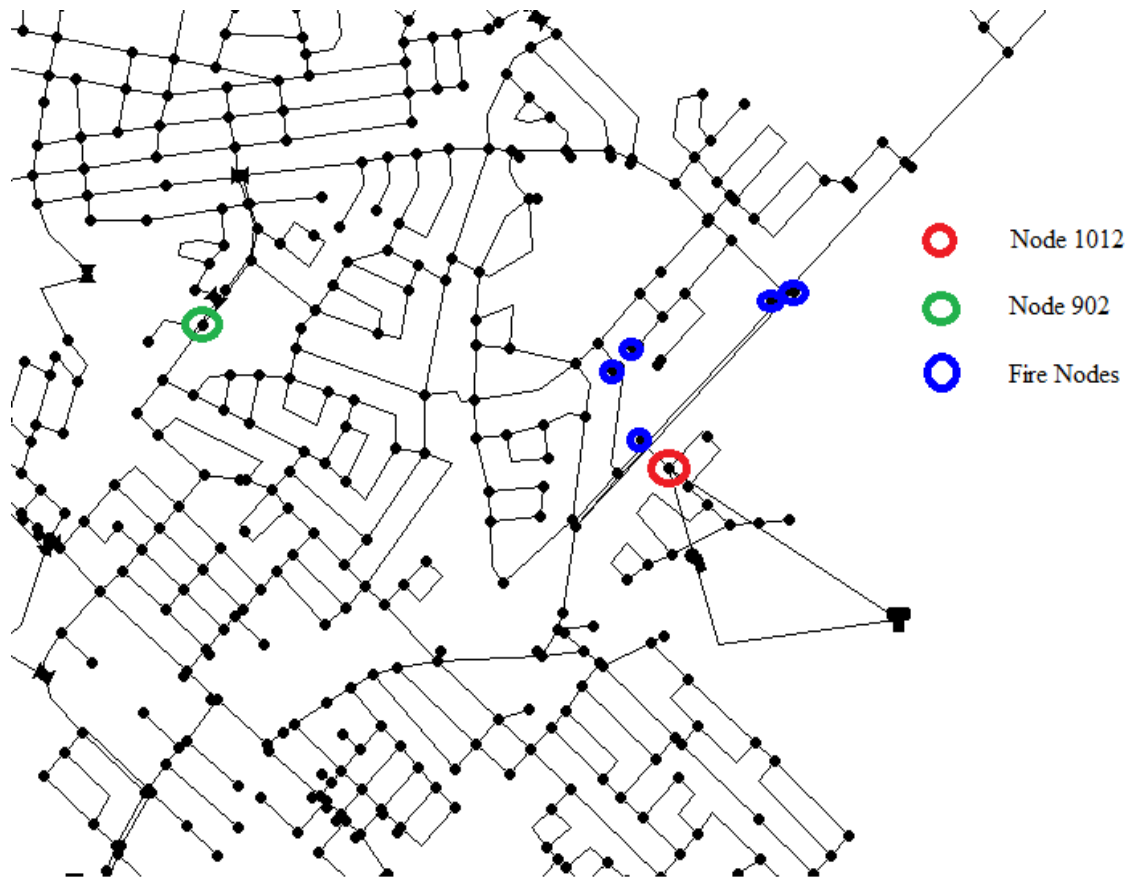


Figure 4.6 : New Tank location and fire flow allocation in the south-east of Saint-Jerome network

It should be noted that Basile (2008) used only the deterministic environment to solve the problem. However, the aim of the study case was not to compare the simulation approach itself. It would not be fair to compare a solution found with a stochastic optimization algorithm with one found with a deterministic approach. The aim was to show the limits of a pseudo-manual optimization approach by trial and error compared to the use of a mathematical optimization method which allows for an adequate exploration of the solution space. It is obvious that the implemented approach does not guarantee the global optimum as is the case for any approach intended to solve these kinds of problems. However, it does provide reasonable solutions during reasonable times to help in meeting this prominent challenge decision-makers face with real-world design problems.

4.2.5 Conclusion

A procedure for the robust design and operation of WDS presented marks the emphasis on the inclusion of water quality as part of the objective functions in the stochastic optimization process, the integration of the correlations in the stochastic nodal demands, the use of a real-life case study including tank design, and the demand for a preliminary analysis aiming to reduce the number and the range of decision variables before the optimization process. The proposed approach was applied to a large and complex real-life network design problem dealing with the introduction of a tank in an existing network. The approach was able to find a more reliable solution in terms of the considered objective functions and especially in the perspective of comparing the level of water quality against results from past studies on the same design problem. It is therefore concluded that the proposed approach possesses a potential in solving real-world WDS design problems. However, given the diversity of constraints which could be required for each network, further applications on various networks may provide insights regarding ways to improve the proposed approach. The process could very well be applicable in future especially on WDSs with more flexible terms of design constraints and can be made to test the other aspects of the proposed approach such as considerations involving pipe design.

CHAPITRE 5 MODÉLISATION DU RÉSEAU

5.1 Introduction

L'objectif du présent chapitre complémentaire aux articles 2 et 3 est de présenter certains détails du modèle développé. Les différentes classes de base utilisées pour conceptualiser le réseau de distribution sont exposées dans un premier temps. Ensuite les étapes clés du programme incluant les différentes fonctions sollicitées sont présentées ainsi que les hypothèses éventuelles et leurs motivations. Un ensemble d'indicatifs sur les éventuelles fonctions ou classes à modifier sont fournis pour orienter dans la personnalisation du modèle pour un cas particulier de réseau.

5.2 Conceptualisation du réseau de distribution

Le modèle numérique présenté dans ce chapitre conceptualise le RDEP à l'aide de plusieurs objets regroupés en deux principaux groupes: les objets physiques et les objets non physiques.

5.2.1 Objets physiques

Comme illustré à la Figure 5.1, les objets physiques sont ceux qui peuvent être vus sur le graphique du modèle numérique du réseau. Quatre types d'objets composent cette catégorie : les nœuds (nœud de demande, bache et réservoir), les arcs (conduite, pompe et vanne), les points de passage et les textes.

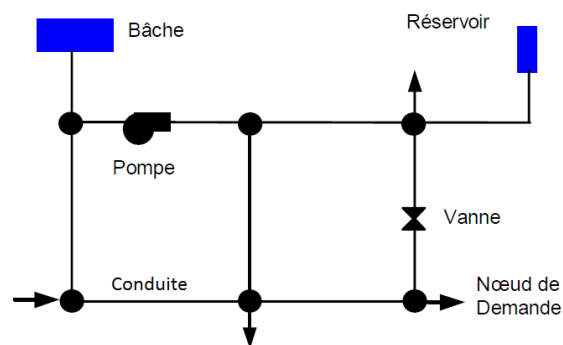


Figure 5.1 : Objets physiques d'un système de distribution d'eau potable (Rossman, 2000)

5.2.1.1 Objet Nœud

Dans le présent modèle numérique du réseau, l'objet nœud est subdivisé en trois sous objets : les nœuds de demande, les bâches et les réservoirs.

5.2.1.1.1 Objet Nœud de demande

Un nœud de demande peut avoir un débit positif, négatif ou nul. Les données requises sur un nœud de demande sont l'altitude, la demande et la qualité initiale de l'eau. L'état des nœuds de demande est caractérisé à chaque pas de temps de simulation par la pression et la qualité de l'eau.

5.2.1.1.2 Objet Bâche

Une bâche peut représenter une source externe de capacité infinie. Les données requises sur une bâche sont la charge totale et la qualité initiale de l'eau.

5.2.1.1.3 Objet Réservoir

Les réservoirs ont une capacité de stockage et le volume d'eau stockée peut varier avec le temps. Les données requises sur un réservoir sont l'altitude du radier, le diamètre, les niveaux initial, minimal et maximal de l'eau et la qualité initiale de l'eau. L'état des réservoirs est caractérisé à chaque pas de temps de simulation par la pression et la qualité de l'eau.

5.2.1.2 Objet Arc

Dans le présent modèle numérique du réseau, l'objet arc est subdivisé en trois sous objets : les conduites, les pompes et les vannes.

5.2.1.2.1 Objet Conduite

Une conduite relie deux nœuds du réseau. Les données requises sur une conduite sont le nœud initial et final, le diamètre, la longueur et le coefficient de rugosité. L'état des conduites est caractérisé à chaque pas de temps de simulation par la vitesse de l'eau.

5.2.1.2.2 Objet Pompe

Une pompe ajoute l'énergie à l'eau et augmente donc sa charge hydraulique. Les données requises sur une pompe sont le nœud d'aspiration et de refoulement, la courbe caractéristique et

la courbe de modulation de la vitesse. L'état des pompes est caractérisé à chaque pas de temps de simulation par le débit et la hauteur d'eau.

5.2.1.2.3 Objet Vanne

Une vanne régule la pression ou le débit en un point donné du réseau. Les données requises sur une vanne sont le nœud d'entrée et de sortie, le diamètre et le consigne de fonctionnement. L'état des vannes est caractérisé à chaque pas de temps de simulation par le débit et la perte de charge.

5.2.1.3 Objet Point de passage

Un point de passage représente ici un point de transition emprunté par une conduite pour aller vers son nœud final. Ces points permettent de refléter la courbure du chemin suivi par la conduite. Les seules données requises pour ces points sont leurs coordonnées qui ne sont pas utilisées lors de la simulation et servent uniquement pour la visualisation graphique du réseau.

5.2.1.4 Objet Texte

Le texte ici permet d'ajouter des étiquettes à certains éléments du réseau pour les décrire. Le texte n'est pas pertinent pour la simulation. Il sert uniquement pour la visualisation graphique du réseau.

5.2.2 Objets non physiques

Les objets non physiques ne peuvent être vus sur le graphique du modèle numérique du réseau. Les objets composant cette catégorie sont : la demande, la courbe de modulation, les courbes, la commande simple, la commande élaborée, l'énergie, la source, la réaction, le temps, le rapport, les options et le fond d'écran.

5.2.2.1 Objet Demande

Cet objet est un supplément aux informations fournies aux nœuds de demande et permet d'attribuer plusieurs catégories de demandes aux nœuds.

5.2.2.2 Objet Courbe de modulation

Une courbe de modulation est composée ici d'un ensemble de multiplicateurs qui peuvent être appliqués à une valeur de base pour lui permettre de varier dans le temps. On peut assigner des

courbes de modulation à la demande d'un nœud, au niveau d'une bêche, à la vitesse de rotation d'une pompe, à la qualité de l'eau dans une source et au prix de l'énergie.

5.2.2.3 Objet Courbe

Une courbe contient des couples de données ayant une relation entre elles. Les types de courbes utilisés dans ce modèle numérique sont : courbe caractéristique, courbe de rendement, courbe de volume et courbe de perte de charge.

- La courbe caractéristique est associée à une pompe et représente le rapport entre la charge et le débit qu'une pompe peut fournir à sa vitesse nominale.
- La courbe de rendement aussi associée à une pompe représente le rapport entre le rendement et le débit pompé. Cette courbe est utilisée pour les calculs d'énergie.
- La courbe de volume est associée à un réservoir et représente le rapport entre le volume d'eau et le niveau d'eau dans le réservoir. Cette courbe est requise lorsque le réservoir n'a pas une forme régulière.
- La courbe de perte de charge est associée à une vanne et représente le rapport entre la perte de charge et le débit traversant la vanne. Cette courbe est requise lorsque les pertes de charge ne suivent pas une formule hydraulique standard.

5.2.2.4 Objet Commande simple

Une commande simple permet de changer l'état ou la consigne de fonctionnement d'un arc selon le niveau d'eau dans un réservoir, la pression à un nœud ou un instant précis donné. Ces commandes sont régulièrement vérifiées et exécutées au cours de la simulation.

5.2.2.5 Objet Commande élaborée

Une commande élaborée permet de déterminer l'état ou la consigne d'un arc basé sur une combinaison de conditions pour le pas de temps en cours.

5.2.2.6 Objet Énergie

Cet objet permet de définir les paramètres utilisés pour calculer la consommation d'énergie et son coût.

5.2.2.7 Objet Source

Cet objet permet de définir la position et les caractéristiques des éventuels points d'injection d'une substance modifiant la qualité de l'eau.

5.2.2.8 Objet Réaction

Cet objet permet de définir les paramètres relatifs aux réactions chimiques qui ont lieu dans le réseau.

5.2.2.9 Objet Temps

Cet objet permet de définir les différents paramètres temporels qui contrôlent le déroulement de la simulation. Ces paramètres sont par exemple la durée totale de la simulation, le pas de temps pour la simulation hydraulique et de la qualité de l'eau ou l'heure de début des simulations.

5.2.2.10 Objet Rapport

Cet objet permet de décrire le contenu d'un rapport des résultats de la simulation. A titre d'exemple, on peut choisir d'inclure ou non le rapport sur la consommation d'énergie.

5.2.2.11 Objet Options

Cet objet permet de définir les différentes options de simulation. On définit dans cet objet les paramètres tels que les unités du débit, l'équation de perte de charge à utiliser ou le paramètre de qualité.

5.2.2.12 Objet Fond d'écran

Cet objet permet d'identifier une image pour l'utiliser comme fond, et définir ses dimensions. Il sert uniquement pour la visualisation graphique du réseau.

5.3 Présentation générale du modèle

Dans cette section, les différentes étapes à suivre pour l'utilisation du modèle ainsi que certaines fonctions pertinentes ou équations utilisées dans le modèle sont présentées.

5.3.1 Préparation des données par l'utilisateur

Cette étape est très importante car elle permet de fournir la configuration du réseau et les différents paramètres propres au réseau. Il faut commencer par mettre en place le fichier d'entrée EPANET associé au réseau qu'on veut dimensionner. Un formulaire type de format csv doit être rempli à cet effet. Pour voir un exemple de fichier d'entrée à mettre en place, il faut se référer à l'Annexe 1.

Compte tenu de certaines difficultés liées à la programmation et vu la nécessité de généraliser, certaines restrictions sont imposées quant aux paramètres choisis à savoir :

- L'identifiant ID de toutes les composantes du réseau doit être un nombre entier.
- L'identifiant ID d'une courbe de modulation doit être un entier non nul.
- La section [REPORT] doit être saisie telle qu'elle afin de s'assurer qu'un rapport complet sera généré par EPANET à chaque simulation.
- Le paramètre de qualité à entrer dans ce fichier est « Age » c'est à dire le temps de séjour de l'eau. C'est cet indicateur de qualité qui est utilisé dans le programme.
- S'il y a des étiquettes de texte dans le réseau, chaque étiquette doit s'écrire sans espace et suivi obligatoirement par l'ID d'un nœud d'ancrage.

Par ailleurs, EPANET impose que l'identificateur ID des composantes représente une clé primaire pour chaque catégorie à savoir :

- Les Nœuds du réseau : nœuds de demande; bâches et réservoirs
- Les Arcs du réseau : conduites; pompes et vannes
- Les courbes (« Curves »)
- Les courbes de modulation (« Patterns »)

Cette contrainte signifie par exemple qu'un nœud de demande et un réservoir doivent avoir des ID différents de même qu'une courbe caractéristique et une courbe de rendement. Par contre, rien n'empêche qu'une conduite et un nœud de demande ou une courbe aient le même identificateur.

Il faut noter que les sections relatives à la configuration graphique telle que les coordonnées ou les points de passage ne sont pas obligatoires pour les calculs. Elles servent juste à visualiser la

configuration du réseau quand EPANET est exécuté en mode graphique. Les valeurs fournies dans ce fichier d'entrée pour les variables d'optimisation ne sont pas prises en compte par le modèle d'optimisation mais elles servent uniquement pour la simulation initiale en vue de confirmer la validité du réseau. Une fois que le fichier csv est prêt, son chemin d'accès est édité dans le programme et il est donc prêt à être lu et enregistré par ce dernier.

5.3.2 Lecture et enregistrement des données par le programme

Des variables d'accueil créées par le programme permettent d'enregistrer toute information pertinente contenue dans le fichier d'entrée et nécessaire au fonctionnement d'EPANET. Pour ce faire, le manuel d'utilisation EPANET a été analysé minutieusement et les variables de base ont été regroupées en classes et sous classes. Les classes de base ont été décrites dans la section 5.2 précédente intitulée « conceptualisation du réseau de distribution ». Pour un projet donné, il y a tout un ensemble d'éléments (objets) de même classe. Ces éléments ont alors été regroupés dans les vecteurs comme montré au Tableau 5.1. Une fois en possession du fichier d'entrée, le programme commence d'abord par compter le nombre d'éléments de chaque vecteur contenu dans ce fichier afin d'initialiser la taille des vecteurs. Les données sont ensuite enregistrées en totalité dans ces vecteurs et d'autres objets simples ou variables. Ces variables sont très importantes et ce sont elles qui vont être manipulées dans tout le processus d'optimisation.

Tableau 5.1 : Vecteurs de classes élémentaires pour le projet

VECTEUR	TYPE CONTENU	TAILLE
Ensemble _{NoeudDemandes}	Nœud de demande	N _{NoeudDemandes}
Ensemble _{Baches}	Bâche	N _{Baches}
Ensemble _{Reservoirs}	Réservoir	N _{Reservoirs}
Ensemble _{Conduites}	Conduite	N _{Conduites}
Ensemble _{Pompes}	Pompe	N _{Pompes}
Ensemble _{Vannes}	Vanne	N _{Vannes}
Ensemble _{CourbeModulations}	Courbe de modulation	N _{CourbeModulations}
Ensemble _{Courbes}	Courbe	N _{Courbes}
Ensemble _{CommandeSimple}	Commande simple	N _{CommandeSimple}
Ensemble _{CommandeElaborees}	Commande élaborée	N _{CommandeElaborees}
Ensemble _{PointPassages}	Point de passage	N _{PointPassages}
Ensemble _{Textes}	Texte	N _{Textes}

A ce stade, le programme connaît le système d'unités (Métrique/Américaine) utilisée dans le projet. C'est ce système qui impose le choix des équations de calcul dans tout le processus d'optimisation. Le système d'unités est quant à lui imposé par l'unité du débit fournit dans le fichier d'entrée. Le Tableau 5.2 donne les unités du débit associées aux deux systèmes d'unités.

Tableau 5.2 : Système d'unités en fonction du débit

	UNITES METRIQUES	UNITES AMERICAINES
	LPS (l/s)	CFS (ft ³ /s)
	LPM (l/min)	GPM (gal/min)
DEBIT	MLD (méga l/j)	MGD (million gal/day)
	CMH (m ³ /h)	IMGD (Imperial MGD)
	CMD (m ³ /j)	AFD (acre foot/day)

5.3.3 Traitement préliminaire des données

Ce traitement permet de se faire une idée sur la répartition des zones de pression dans le réseau sur la base de la topographie fournie. A cette étape du processus, on se fait également une idée sur le diamètre de référence (Équations 3.1 ou 4.1).

5.3.3.1 Calcul des zones de pression.

Une fois que les données du réseau sont enregistrées par le programme, une compilation est faite afin d'estimer le nombre de zones de pression qu'on peut distinguer dans le réseau. Ce nombre est évalué afin de déterminer le nombre maximal de réservoirs à proposer. Pour ce faire, les calculs élémentaires suivants sont effectués :

- Élévation minimale E_{\min} du réseau
- Élévation maximale E_{\max} du réseau
- Élévation moyenne E_{moy} du réseau
- Nombre de Zones de Pression N_{Zones} (entier)
- Altitudes minimaux A_{\min} des radiers de réservoirs selon la zone
- Altitudes maximaux A_{\max} des radiers de réservoirs selon la zone
- Pourcentage P_{Zone} de clients dans chaque zone de pression

Le calcul de ces paramètres est présenté dans les articles 2 et 3. Les équations utilisées sont automatiquement ajustées au système d'unités.

$$\text{Si } P_{Zone} < \frac{1}{3} \frac{100}{N_{Zones}} \Rightarrow \text{Groupe de points isolés} \quad (5.1)$$

La valeur d'altitude minimale de radier du réservoir est évaluée de telle sorte les consommateurs reçoivent au moins la pression minimale. De la même manière, la valeur d'altitude maximale de radier du réservoir est évaluée de telle sorte les consommateurs reçoivent au plus la pression maximale. Ces valeurs ont pour but de borner l'espace de recherche des solutions dans le processus d'optimisation.

Dans cette thèse, il est suggéré de prévoir au plus un réservoir par zone représentative (zone non isolée).

5.3.3.2 Calcul du diamètre de référence

Le diamètre de référence est abordé dans les articles 2 et 3 (Équations 3.1 ou 4.1) mais avec un débit en m³/s. Si le débit est dans une autre unité, il est multiplié par un coefficient α selon l'unité tel que présenté au Tableau 5.3.

Tableau 5.3: Coefficient de conversion du débit en m³/s

UNITES METRIQUES		UNITES AMERICAINES	
Unité	α	Unité	α
LPS	0.001	CFS	0.028316847
LPM	0.0000167	GPM	0.0000631
MLD	0.011574074	MGD	0.043819444
CMH	0.000277778	IMGD	0.052629641
CMD	0.000011574	AFD	0.0142764102

Le raisonnement utilisé pour calculer ce diamètre est : si on se base exclusivement sur le critère de coût dans l'optimisation, le diamètre maximal du réseau optimisé ne devrait pas dépasser ce diamètre de référence.

5.3.4 Simulation EPANET et enregistrement des résultats

Afin de préparer le modèle à une simulation EPANET, le programme écrit les données enregistrées dans le fichier d'entrée EPANET. Il écrit également une copie en format texte afin de s'assurer (au besoin) manuellement que les données ont été transcrites de façon fidèle. Une série de commandes dos est introduite pour lancer la simulation EPANET et sauvegarder les résultats de cette simulation dans un fichier texte. Le rapport sauvegardé est un rapport complet tel que spécifié dans le fichier d'entrée. Les résultats enregistrés à chaque pas de temps pour les différents paramètres variables sont :

- Nœuds de demande : demande actuelle; charge totale; pression; qualité
- Bâches : débit d'entrée; charge totale; pression; qualité
- Réservoirs : débit d'entrée; altitude de la surface; niveau; qualité
- Conduites : débit; vitesse; perte de charge unitaire; état actuel
- Pompes : débit; qualité; hauteur; état actuel
- Vannes : débit; vitesse; perte de charge; état actuel

Le programme EPANET enregistre également la consommation journalière d'énergie et pour cela, il faut s'assurer d'avoir entré un coût d'énergie dans le fichier d'entrée.

5.3.5 Traitement des résultats et évaluation de la solution

Ce traitement vise à déterminer le degré de qualité de la solution en cours notamment par rapport au respect des contraintes hydrauliques et aux valeurs prises par les fonctions objectives. Le respect des contraintes, les fonctions objectives ainsi que le critère de classement des solutions sont discutés dans les articles 2 et 3 en fonction de l'environnement de travail (déterministe ou stochastique).

5.3.6 Modélisation des variables du problème

Les éléments du réseau considérés affectables par la variabilité sont:

- Les conduites
- Les pompes
- Les réservoirs
- Les vannes

Les variables sont décidées par le concepteur selon le réseau à dimensionner. Ces variables sont discutées dans les articles 2 et 3. Une fois les variables choisies, la forme d'une solution type doit être ajoutée dans le code.

5.3.7 Codage et interprétation des chaînes de solution

La méthode de résolution choisie pour traiter le problème d'optimisation posé est le recours aux algorithmes génétiques. Ces algorithmes évolutionnaires se sont révélés efficaces pour traiter le problème de conception des RDEP. La section suivante présente les classes de base utilisées dans la librairie des algorithmes génétiques.

5.3.7.1 Classes de base de la bibliothèque GALib (Genetic Algorithm Library)

GALib est une bibliothèque C++ d'objets d'algorithmes génétiques. Cette bibliothèque contient des outils pour l'implémentation des algorithmes génétiques dans l'optimisation d'un problème donné, en utilisant une représentation donnée et des opérateurs génétiques selon le problème. La bibliothèque disponible en ligne est codée et maintenu par Wall (1996). L'utilisation de cette librairie fait nécessairement appel à deux classes : un génome et un algorithme génétique. Chaque instance du génome représente une solution élémentaire au problème. Le type d'algorithme génétique utilisé définit comment l'évolution doit se faire. L'algorithme génétique utilise une fonction objective définie par l'utilisateur de la librairie pour déterminer l'adaptation ou la qualité de chaque génome, ce qui détermine sa chance de survie.

Trois décisions sont à prendre pour résoudre un problème en utilisant un algorithme génétique :

- Définir une représentation : la structure du génome
- Définir les opérateurs génétiques : initialisation, mutation, et crossover...
- Définir la fonction objective : évaluation du génome

Les deux premières décisions peuvent être prises en utilisant au maximum les outils proposés par GALib tandis que la troisième est spécifique au problème et laissée aux soins de l'utilisateur.

L'algorithme génétique suit essentiellement le processus de la Figure 5.2. Il faut spécifier le critère d'arrêt à l'algorithme. Le critère d'arrêt typiquement utilisé est le nombre de générations,

mais on peut utiliser aussi la qualité de la meilleure solution, la convergence de la population ou un autre critère spécifique au problème.

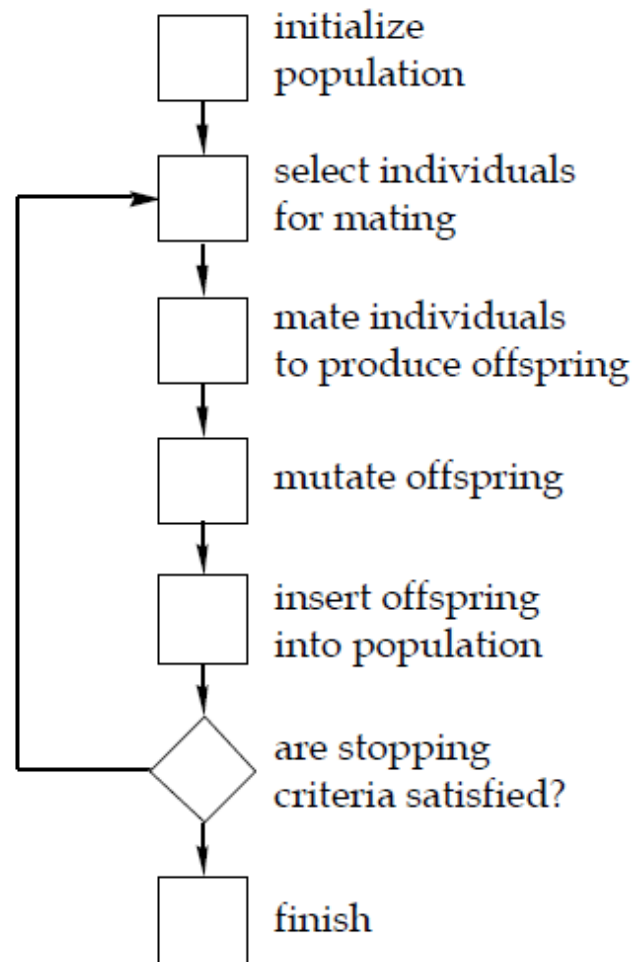


Figure 5.2 : Principe de base de l'Algorithme génétique

La librairie propose quatre types d'algorithmes génétiques :

- Simple Genetic Algorithm: les populations ne chevauchent pas
- Steady-State Genetic Algorithm: chevauchement des populations
- Incremental Genetic Algorithm: générations composées de 1 ou 2 enfants seulement
- Deme Genetic Algorithm: plusieurs populations évoluent en parallèle avec possibilité de migration d'une population à une autre

Les classes et leur hiérarchie se présentent comme suit :

GA GeneticAlgorithm

- GASteadyStateGA (overlapping populations)
- GASimpleGA (non-overlapping populations)
- GAIncrementalGA (overlapping with custom replacement)
- GADemeGA (parallel populations with migration)

GAStatistics

GAParameterList

GAPopulation

GA ScalingScheme

- GANoScaling
- GALinearScaling
- GA SigmaTruncationScaling
- GAPowerLawScaling
- GASharing

GA SelectionScheme

- GARankSelector
- GARouletteWheelSelector
- GATournamentSelector
- GAUniformSelector
- GA SRSSelector
- GADSSelector

GA Genome

- GA1DBinaryStringGenome
 - GABin2DecGenome
- GA2DBinaryStringGenome
- GA3DBinaryStringGenome
- GA1DArrayGenome<>
 - GA1DArrayAlleleGenome<>
 - GAStringGenome (same as GA1DArrayAlleleGenome<char>)

```

    GRealGenome (same as GA1DArrayAlleleGenome<float>)
GA2DArrayGenome<>
    GA2DArrayAlleleGenome<>
GA3DArrayGenome<>
    GA3DArrayAlleleGenome<>
GATreeGenome<>
GAListGenome<>

```

```

GAArray<>          GAAlleleSetArray<>  GAAlleleSet<>
GABinaryString     GABin2DecPhenotype
GATree<>           GATreeIter<>
GAList<>           GAListIter<>

```

5.3.7.2 Implémentation de la bibliothèque GALib au projet de recherche

Les décisions prises pour permettre l'utilisation de GALib dans le modèle proposé ont été abordées dans l'article 2. Il s'agit notamment de la taille de la population, le nombre de générations, la probabilité de mutation et la probabilité de crossover. L'algorithme génétique en régime stable (GASteadyStateGA) a été utilisé. Ce choix a été fait pour accélérer le processus de convergence (chevauchement des populations) sans pour autant manipuler trop de génomes vu l'importance des simulations hydrauliques et de qualité de l'eau à effectuer sur chaque génome.

5.3.8 Liste de quelques fonctions et leurs rôles dans le modèle

Dans cette section, certaines fonctions implémentées dans le modèle ainsi que leurs rôles sont présentées. L'appel à l'une ou l'autre de ces fonctions peut dépendre de l'environnement retenu pour la conception du réseau à savoir déterministe ou stochastique.

5.3.8.1 Échantillonnage avec indépendance

Cette fonction est utilisée dans le volet stochastique. Elle permet de générer une série d'échantillons représentant les demandes aléatoires aux nœuds. Ces demandes aux nœuds sont supposées indépendantes entre elles. Plusieurs méthodes d'échantillonnage peuvent être utilisées à cette fin. Dans le présent modèle, la méthode d'échantillonnage « Latin Hypercube » est utilisée pour des raisons évoquées dans l'article 3. Cette méthode échantillonne en fait une matrice de probabilités de non dépassement qui sont converties en débits suivant une loi de

distribution. Chaque ligne dans cette matrice correspond à l'ensemble des demandes aléatoires aux nœuds pour une simulation. Le nombre de lignes ou taille de l'échantillon correspond au nombre de simulations stochastiques considéré.

5.3.8.2 Échantillonnage avec corrélation

Cette fonction est utilisée dans le volet stochastique au cas où on décide d'intégrer une corrélation dans les demandes aux nœuds. La motivation pour la considération de cet aspect vient du fait que les demandes aux nœuds peuvent être dépendantes dans une certaine mesure de facteurs incertains qui affectent le système dans son ensemble tels que la température. Ces facteurs tendraient à augmenter ou diminuer simultanément la demande à tous les nœuds. Pour faire cet échantillonnage, le modèle a besoin d'une matrice de corrélations cibles ainsi que la matrice d'échantillonnage à transformer. La matrice d'échantillonnage est générée par la fonction précédente. Pour ajouter des corrélations, les étapes utilisées par la présente fonction sont : (1) calcul de la matrice de corrélations actuelles; (2) Décomposition de Cholesky des matrices de corrélations actuelles et cibles; (3) Calcul d'une matrice intermédiaire dont les corrélations des rangs sont proches de celles souhaitées; (4) Réarrangement des éléments de la matrice d'échantillonnage de telle sorte que les rangs dans cette matrice coïncident avec les rangs de la matrice intermédiaire.

5.3.8.3 Conversion Diamètre Économique

Cette fonction permet de convertir un nombre en diamètre commercial correspondant. L'ensemble des diamètres commerciaux est classé et numéroté. À chaque numéro est donc associé un et un seul diamètre commercial.

5.3.8.4 Redondance

Cette fonction détermine pour chaque nœud la liste des conduites qui lui sont connectées. Cet inventaire permettra au modèle d'intégrer l'uniformité des diamètres à chaque nœud dans les calculs de performances. Il est souhaité que les conduites connectées à un même nœud ne soient pas très variables en diamètres.

5.3.8.5 Résilience du réseau

Cette fonction peut être utilisée dans le volet déterministe comme un des objectifs de la conception. Il est présenté dans l'article 2. Il s'agit de l'indice de résilience modifié pour intégrer des pondérations relatives à la redondance évoquée dans la fonction précédente.

5.3.8.6 Temps de séjour déterministe

Cette fonction permet de calculer le temps de séjour en environnement déterministe tel que présenté dans l'article 2. Pour chaque configuration du réseau, un seul scénario de demandes aux nœuds est considéré pour la simulation de longue durée. Le temps de séjour résultant est donc unique.

5.3.8.7 Temps de séjour stochastique

Cette fonction permet de calculer le temps de séjour en environnement stochastique tel que présenté dans l'article 3. Pour chaque configuration du réseau, plusieurs scénarios (échantillons) de demandes aux nœuds sont considérés pour les simulations stochastiques de longue durée. Le temps de séjour stochastique est donc évalué comme étant la moyenne des temps de séjour déterministes de chaque échantillon.

5.3.8.8 Coût total

Cette fonction permet de calculer le coût total qui est composé du coût d'investissement et celui d'opération. La valeur du coût total dépend si on est en environnement déterministe ou stochastique. Le coût journalier d'énergie par exemple dépend des demandes aux nœuds. Dans le volet stochastique, une moyenne est considérée pour ce coût sur l'ensemble des échantillons simulés.

5.3.8.9 Pénalité déterministe

Cette fonction permet de calculer la pénalité relative au respect des contraintes notamment celles sur les vitesses, les pressions ou les niveaux des réservoirs selon les contraintes locales du projet.

5.3.8.10 Pénalité stochastique

Cette fonction est utilisée en environnement stochastique et joue le même rôle que la précédente. Les valeurs devant respecter les contraintes étant variables selon l'échantillon simulé, la valeur de la pénalité dans ce cas reflète une pénalité moyenne.

5.3.8.11 Rang Pareto déterministe

Ce rang est un classement des solutions dans une population selon les valeurs des fonctions objectives associées à ces solutions. Les fonctions objectives utilisées pour ce classement sont présentées dans l'article 2.

5.3.8.12 Rang Pareto stochastique

Cette fonction joue le même rôle que la précédente. Les fonctions objectives utilisées dans le volet stochastique sont présentées dans l'article 3.

5.3.8.13 Évaluation déterministe

Cette fonction permet d'évaluer une solution et lui assigner les valeurs pour chaque fonction objective considérée dans le volet déterministe. L'évaluation se fait selon les étapes suivantes :

- Réception de la solution codée de la part de l'algorithme génétique.
- Décodage de la solution pour refléter les vraies valeurs des variables. Ce décodage fait par exemple appel à la fonction de conversion du diamètre économique abordée précédemment.
- Écriture des paramètres décodés de la solution dans le fichier d'entrée EPANET.
- Simulation hydraulique et de qualité de l'eau de la solution par le modèle de simulation EPANET.
- Enregistrement des résultats de la simulation et évaluation des fonctions objectives considérées en faisant appel à des fonctions abordées précédemment et dans l'article 2.

5.3.8.14 Évaluation stochastique

Cette fonction permet d'évaluer une solution et lui assigner les valeurs pour chaque fonction objective considérée dans le volet stochastique. L'évaluation se fait selon les étapes suivantes :

- Réception de la solution codée de la part de l'algorithme génétique.
- Décodage de la solution pour refléter les vraies valeurs des variables.
- Réception d'un échantillon déterministe issu de l'échantillonnage stochastique pour les demandes aux nœuds.
- Écriture des paramètres décodés la solution ainsi que l'échantillon des demandes aux nœuds dans le fichier d'entrée EPANET.
- Simulation hydraulique et de qualité de l'eau de la solution par le modèle de simulation EPANET.
- Enregistrement des résultats de la simulation.
- Répétition des quatre précédentes étapes jusqu'à la fin des échantillons
- Évaluation des fonctions objectives considérées en faisant appel à des fonctions abordées précédemment et dans l'article 3.

Les algorithmes complets pour les processus d'optimisation sont présentés dans les articles 2 et 3. La section suivante présente les fonctions pouvant être personnalisées pour la conception optimale d'un réseau dont les exigences et les données sont différentes de celles actuellement implémentées dans le modèle développé.

5.4 Personnalisation du modèle

Plusieurs changements peuvent être apportés au code pour l'adapter à un cas particulier de réseau. En fonction des données à considérer, les changements peuvent être plus ou moins longs. Les différentes personnalisations envisageables sont présentées dans les sections qui suivent. Il est à noter que la liste n'est pas exhaustive.

5.4.1 Diamètres commerciaux

La liste des diamètres commerciaux est facile à personnaliser dans le code et implique une modification mineure. Il suffit de modifier la fonction de conversion du diamètre économique. Le nombre et les valeurs de ces diamètres peuvent alors changer. Par contre vu que l'algorithme génétique est configurée pour coder la variable du diamètre de la conduite comme un entier

représentant le rang du diamètre dans la liste classée, il n'ya donc pas d'impact sur ce codage. Par contre, la plage de variation de cette variable va changer pour refléter le nombre de diamètres commerciaux disponibles.

5.4.2 Variables

Plusieurs contraintes pourraient s'imposer quant aux paramètres variables du réseau. Ces contraintes sont par exemple :

- Les diamètres de toutes les conduites sont fixes : dans ce cas, il faut sauter l'étape des conduites : (1) dans la construction du génome type, (2) lors du décodage du génome dans les fonctions d'évaluation ainsi que (3) dans l'évaluation du coût total. La fonction de redondance va être évaluée une seule fois au début du processus. Cette fonction doit être alors désactivée dans l'évaluation du génome. Les modifications conséquentes sont mineures.
- Les diamètres des nouvelles conduites sont variables tandis que ceux des conduites existantes sont fixes (extension du réseau): cette configuration demande plus de modifications que la précédente. Pour gérer efficacement cette situation, il faut lors de la préparation des données (pour la mise en place du fichier EPANET initial) regrouper les conduites existantes d'une part et les nouvelles conduites d'autre part. De cette manière, les opérations qui s'appliquaient aux conduites s'appliqueront au lot des nouvelles conduites uniquement. Le décodage du génome et la mise à jour des diamètres va légèrement changer. Par exemple, s'il y'a 12 conduites existantes suivies de 18 nouvelles à installer, la boucle de modification des diamètres de conduite ira de 13 jusqu'au nombre maximal de diamètres commerciaux considérés. Le même raisonnement s'applique à l'évaluation du coût total.
- Nettoyage des conduites existantes : dans le décodage de la solution, la rugosité de la conduite sera remplacée pour une conduite nettoyée. En cas de nettoyage, la conduite devient plus lisse. La fonction de coût sera modifiée pour refléter le coût de la réhabilitation des conduites.

- La politique d'opération des pompes est fixe : cette politique est alors entrée dans le fichier EPANET initial. Il faut alors sauter l'étape d'opération des pompes dans la construction du génome type et lors de son décodage dans les fonctions d'évaluation.
- Les emplacements des réservoirs sont imposés : si tous les réservoirs ont des emplacements imposés, alors il faut simplement sauter l'étape des emplacements dans la construction du génome type et lors de son décodage dans les fonctions d'évaluation. Si c'est seulement pour certains des réservoirs, il faut faire une séparation lors de la préparation du fichier d'entrée initial et appliquer la même philosophie de modification du code que pour le cas de l'extension du réseau évoquée précédemment pour les conduites.

5.4.3 Contraintes

La personnalisation des contraintes a beaucoup plus d'impact sur le code que celle des variables. Actuellement, la pression minimale requise aux nœuds de demande est considérée constante. Il en est de même pour la pression maximale aux nœuds, la vitesse minimale et maximale dans les conduites. Si ces valeurs doivent être personnalisées pour chaque nœud et pour chaque conduite, des modifications majeures seront requises dans le code telles que :

- Ajout d'une fonction pour la lecture séparée d'un fichier contenant les valeurs des contraintes pour chaque nœud et chaque conduite. Cela nécessite également la création de vecteurs pour sauvegarder ces contraintes. La taille des vecteurs pour les contraintes aux nœuds (respectivement sur les conduites) sera égale au nombre de nœuds (respectivement au nombre de conduites).
- Modification de la fonction de calcul de la Résilience du réseau ainsi que les fonctions de pénalité.

5.4.4 Fonctions objectives

La personnalisation des fonctions objectives telles que le coût ou la considération d'objectifs divers peut aussi donner lieu à des modifications importantes dans le code.

5.4.4.1 Coût

Il faut personnaliser le calcul des coûts dans la procédure de coût total. Ceci pourrait aussi nécessiter d'ajouter des attributs aux classes de base. Par exemple, si plusieurs matériaux sont considérés pour les conduites, le prix étant fonction du matériau, on pourrait ajouter un attribut « type » dans la définition de la classe conduite. De cette façon, dans l'implémentation de la fonction de coût total, cet attribut sera vérifié avant d'attribuer le bon coût.

5.4.4.2 Autres fonctions objectives

Plusieurs objectifs alternatifs peuvent être envisagés dans la conception optimale des réseaux de distribution d'eau potable. Parmi ces alternatives, on peut citer :

- Minimiser l'indice de déficit hydraulique : défaut de débit fournit ou baisse de niveau des réservoirs par exemple.
- Minimiser le déficit en débit d'incendie : le débit et la pression doivent être suffisants pour combattre un incendie.
- Minimiser les dommages annuels attendus : les faibles pressions peuvent causer par exemple une interruption de la production industrielle tandis que les pressions excessives favorisent les bris de conduites.

La considération d'un indicateur de performance non implémenté dans le code nécessite l'ajout d'une procédure dans le programme. Il faudra modifier également la fonction qui évalue le Rang Pareto pour mettre à jour les objectifs à considérer pour la comparaison des solutions.

5.4.5 Extension des scénarios de demande pour la simulation déterministe

Au lieu de considérer un seul scénario de demandes avec une courbe de modulation horaire, il peut s'agir de considérer plusieurs scénarios indépendants, chacun avec ses contraintes propres. Comme exemples de scénario, on peut avoir : demande de la journée moyenne, demande de pointe instantanée ou une combinaison d'incendies. Pour une telle situation, beaucoup de modifications seront requises. Il faut : créer de nouveaux vecteurs dans le programme pour enregistrer les différents scénarios; créer une fonction de lecture des scénarios; modifier la fonction d'évaluation d'une solution en prenant par exemple des valeurs moyennes (entre les

différents scenarios) pour les indicateurs hydrauliques de performance; modifier le calcul de la pénalité.

5.4.6 Paramètres de l'algorithme génétique

La modification de ces paramètres a en général très peu d'impact sur les fonctions du code. L'appel aux paramètres suivants est très local dans le code: taille de la population, nombre de générations, probabilité de mutation et probabilité de croisement. Pour personnaliser ces paramètres, il suffit de modifier leurs valeurs dans le code. Par contre, si on modifie la représentation du génome, il faut modifier également le décodage du génome dans l'évaluation de la solution. On peut choisir également un autre type d'algorithme génétique parmi les types évoqués dans la hiérarchie des classes de l'algorithme génétique. D'autres types externes d'algorithmes génétiques tels que NSGAII ou SPEAII ont été testés mais se sont révélés plus longs en termes de durée de simulation et quand le nombre de simulation était un peu grand, des erreurs de manque de mémoire ont été rencontrées avec ces algorithmes.

5.4.7 Paramètres d'échantillonnage stochastique

La fonction d'échantillonnage est indépendante du programme d'autant plus qu'elle permet de fournir un échantillon de demandes aléatoires au programme pour la simulation, peu importe la méthode d'échantillonnage utilisée. Trois versions d'échantillonnage « Latin Hypercube » sont disponibles dans le code pour être utilisées : le type centré, extrême et aléatoire. Le type actuellement utilisé est le centré mais les deux autres peuvent être également utilisés et la différence n'est pas trop grande. Pour utiliser d'autres types externes, il faut les programmer ou les importer. Pour la loi de distribution, d'autres fonctions de probabilités peuvent être utilisées telles que la loi normale. Il faut dans ce cas modifier la formule de conversion de la probabilité en demande.

5.4.8 Modèle de simulation

Le modèle de simulation EPANET est celui qui est utilisé dans le code. Il s'agit de la version classique d'EPANET qui traite les demandes comme satisfaites et calcule les pressions aux nœuds, celle-ci pouvant être négatives. Il existe une version modifiée (extension) d'EPANET dans la littérature qui délivre un débit au nœud en fonction de la pression résultante (Morley &

Tricarico, 2008). Par contre, cette extension n'est pas disponible gratuitement en ligne. En fonction de l'approche qu'on souhaite donner à la simulation et des objectifs souhaités pour la conception optimale du réseau, cette version pourrait être mieux adaptée. Si cette extension d'EPANET devrait être utilisée dans le code en remplacement de la version classique, des modifications doivent être apportées au code : de nouvelles variables s'ajoutent telles que les pressions minimales et souhaitables aux nœuds. Il faut alors modifier les fonctions de lecture et d'écriture des variables; modifier certaines classes de base telles que les attributs des nœuds; implémenter le calcul des fonctions objectives qui vont avec cette approche de calcul.

5.5 Conclusion

Les classes de base utilisées dans la conceptualisation du réseau ont été présentées. Les interactions entre ces classes ainsi que les hypothèses faites ont été discutées et des suggestions ont été faites sur les éventuelles fonctions ou classes à modifier dans un cas particulier.

CHAPITRE 6 COMPARAISON DES APPROCHES SUR UN EXEMPLE D'APPLICATION

6.1 Introduction

L'objectif du présent chapitre est d'appliquer le modèle sur un exemple de réseau en utilisant les différentes approches de conception en vue d'une comparaison. La première section est dédiée à la description du réseau utilisé tandis que la seconde présente les résultats de l'optimisation du réseau par les différentes approches étudiées dans cette thèse.

6.2 Description du système de distribution

Le réseau considéré est un réseau à quatre boucles tiré de Lansey et Mays (1989). Il contient les composantes typiques d'un réseau de distribution d'eau potable (conduite, pompe et réservoir) dont les caractéristiques sont susceptibles de varier dans le présent modèle.

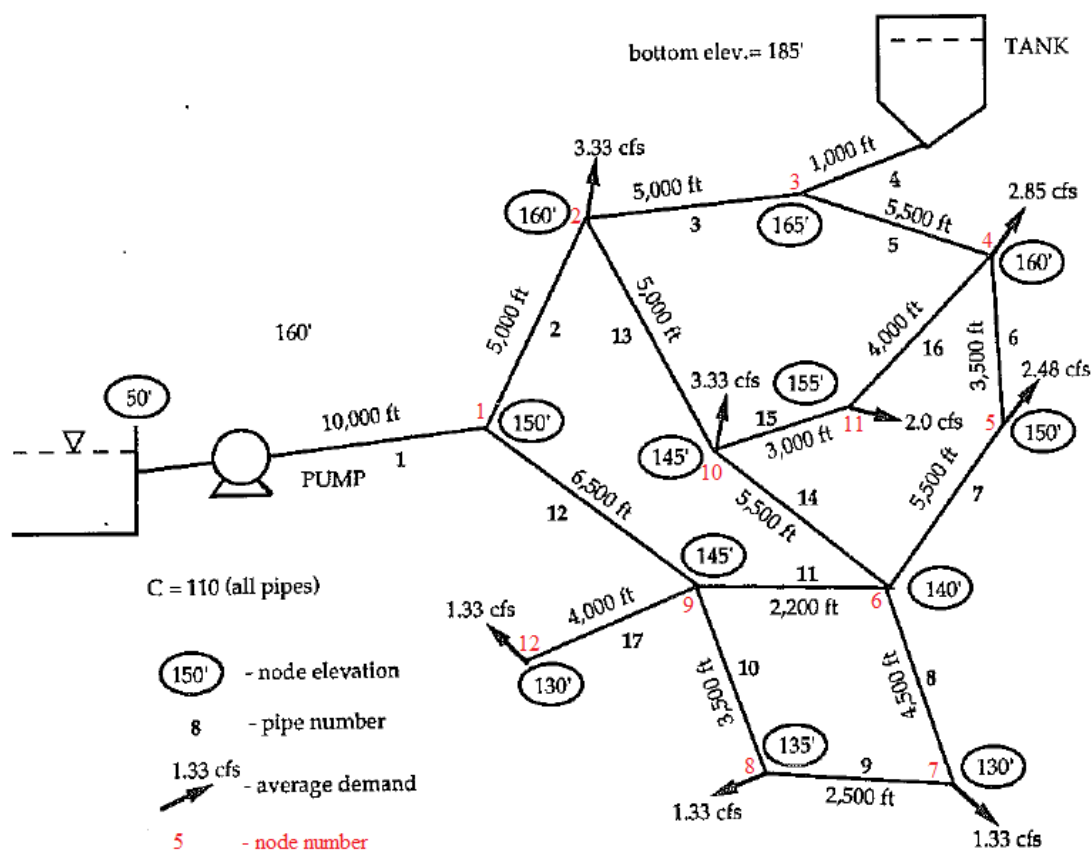


Figure 6.1 : Réseau de l'Exemple (Lansey & Mays, 1989)

Tableau 6.1 : Diamètres commerciaux disponibles

N°	D(mm)	D(in)
1	20	0,75
2	25	1
3	32	1,25
4	37	1,5
5	50	2
6	62,5	2,5
7	75	3
8	100	4
9	125	5
10	150	6
11	200	8
12	250	10
13	300	12
14	350	14
15	400	16
16	450	18
17	500	20
18	600	24
19	750	30
20	900	36
21	1050	42
22	1200	48
23	1350	54
24	1500	60
25	1600	64
26	1650	66
27	1800	72
28	2100	84
29	2400	96
30	2700	108

Ce réseau a été initialement présenté dans l'exemple 5A du manuel d'utilisateur de KYPIPE (Wood, 1980) en tant que réseau typique d'une petite municipalité. Le réseau (sans les vannes) avec la demande moyenne à chaque nœud est présenté à la Figure 6.1. Les diamètres commerciaux disponibles sont présentés au Tableau 6.1.

Une journée typique a été subdivisée en quatre sous-périodes de six heures chacune avec les multiplicateurs de la demande moyenne suivant la courbe de modulation 0,67 – 1,2 – 0,78 – et 1,4. La pression minimale exigée pour ce réseau est de 275.8 kPa (40 psi) pour tous les nœuds. Le coût d'énergie est calculé en supposant un prix de 0,12\$/kWh, un taux d'intérêt de 12%, une période d'amortissement de 20 ans, et un rendement constant de 0,8 pour la pompe.

Le coût total de ce réseau est calculé comme étant la somme de quatre composantes : (1) coût d'investissement des conduites ; (2) coût d'investissement du réservoir ; (3) coût d'investissement de la pompe ; et (4) coût d'opération de la pompe.

Le coût d'investissement C_c d'une conduite est calculé en fonction de sa longueur L et son diamètre D comme à l'équation (6.1).

$$C_c = 1,1152 * L * D^{1,3276} \quad (6.1)$$

Le coût d'investissement C_r du réservoir est calculé en fonction de son diamètre D_r et sa hauteur H_r comme à l'équation (6.2).

$$C_r = 614,32 * D_r^{1,1606} * H_r^{0,5803} \quad (6.2)$$

Le coût d'investissement C_p d'une pompe est calculé en fonction de son débit Q_p et sa hauteur manométrique H_p comme à l'équation (6.3).

$$C_p = 35927,58 * Q_p^{0,7} * H_p^{0,4} \quad (6.3)$$

Afin de tenir compte du remplacement de la pompe, la durée de vie de la pompe a été fixée à 10 ans contrairement à la durée de vie du système qui est de 20 ans. La pompe sera alors remplacée une fois au cours des 20 ans d'opération du système. Cet investissement futur doit être actualisé et le coût total actualisé de la pompe est alors donné par l'équation (6.4) dans laquelle i_r est le taux d'intérêt.

$$C_{p,\text{actualisé}} = C_p * \left(1 + \frac{1}{(1+i_r)^9}\right) \quad (6.4)$$

Le coût d'opération C_{op} de la pompe est défini comme le coût actualisé de la consommation d'énergie sur la durée de vie du système (20 ans). Le coût journalier d'énergie $C_{journalier, \text{énergie}}$ est calculé par *EPANET* en utilisant le prix par kilowattheure fourni. Ce coût est alors ramené sur une base annuelle et actualisé comme à l'équation (6.5).

$$C_{op} = 365 * C_{journalier, \text{énergie}} * \sum_{j=0}^{19} \frac{1}{(1+i_r)^j} \quad (6.5)$$

6.3 Optimisation du système de distribution

Trois approches différentes sont utilisées pour résoudre ce problème afin de comparer les résultats : l'approche déterministe, l'approche stochastique sans les corrélations entre les demandes aux nœuds et l'approche stochastique tenant compte des corrélations entre les demandes aux nœuds.

6.3.1 Approche déterministe

Dans cette approche de résolution du problème, les demandes aux nœuds sont supposées constantes et varient seulement suivant la courbe de modulation journalière fournie dans la description du problème. Afin de comparer les résultats de ce volet déterministe avec ceux de Lansey et Mays (1989), le débit de conception de la pompe a été fixé à 29,33 pi³/s avec une hauteur manométrique de 187 pieds. Une politique d'opération horaire est retenue pour la pompe. Tout au long des 24h de la journée type, l'optimisation doit déterminer l'état (marche/arrêt) à donner à la pompe à chaque heure.

Le traitement des élévations aux nœuds donne une seule zone de pression pour ce réseau. Par conséquent, l'approche proposée dans cette thèse suggère un réservoir au maximum, c'est-à-dire zéro ou un réservoir. Deux alternatives sont alors étudiées pour ce problème : la première avec un réservoir incorporé au réseau et la seconde excluant le réservoir. Aucune réserve d'incendie n'est prévue pour ce réseau.

Alternative 1

Les variables considérées sont : (1) une variable diamètre pour chacune des 17 conduites : ce diamètre est choisi dans la liste des 30 diamètres commerciaux disponibles (voir Tableau 6.1). Avec ces 30 alternatives, le diamètre est modélisé comme un entier prenant des valeurs entre 1 et 30. (2) Une variable état de la pompe pour chacun des 24 pas de temps de simulation. L'état de

la pompe est modélisé comme un entier binaire. (3) Une variable pour l'emplacement du réservoir : cet emplacement est choisit parmi les 12 nœuds du réseau. Avec ces 12 alternatives, l'emplacement est modélisé comme un entier prenant des valeurs entre 1 et 12. (4) Une variable pour le diamètre du réservoir. Cette variable est modélisée comme un réel prenant des valeurs entre 16 pieds et 130 pieds. Un ratio diamètre/hauteur de 0,8 est adopté pour ce réservoir, incluant une revanche de 10% du volume total. Avec cette hypothèse, le diamètre permet de fixer le niveau maximum d'opération du réservoir. (5) Une variable pour l'altitude du radier de réservoir. Cette variable est modélisée comme un réel prenant des valeurs entre 215 pieds et 300 pieds conformément aux Équations (3.3) et (3.4).

Le nombre de variables est donc : 17 pour les conduites, 24 pour la pompe et 3 pour le réservoir ; ce qui donne un total de 44 variables pour le problème d'optimisation.

Les trois fonctions objectives considérées sont : minimiser le coût total, minimiser le temps de séjour maximal dans le réseau et maximiser l'indice de résilience du réseau. Les contraintes qui donnent lieu à une pénalité sont : (1) la pression minimale (275.8 kPa) aux nœuds; (2) la vitesse maximale (3 m/s) dans les conduites et (3) le décalage entre le niveau maximal fixé pour le réservoir et le niveau maximal réel d'opération (résultant de la simulation).

Afin de déterminer les caractéristiques du compromis entre les trois fonctions objectives, l'algorithme génétique a été exécuté avec une taille de population de 50 solutions et un nombre de générations égal à 2000. La frontière de Pareto identifiée par l'algorithme génétique à la fin des itérations contenait 39 solutions. Il est à noter que la résolution de ce problème d'optimisation à trois objectifs résout automatiquement 6 petits sous-problèmes en même temps : 3 problèmes à une fonction objective et 3 problèmes à deux fonctions objectives. Autrement dit, la frontière de Pareto obtenue pour ce problème d'optimisation à trois fonctions objectives contient tous les compromis pour les 6 sous-problèmes. L'exploration et l'analyse des frontières de Pareto des trois problèmes à deux fonctions objectives permettent d'identifier des solutions satisfaisantes.

Les Figures 6.2 à 6.4 montrent la projection de la frontière Pareto dans les trois plans formés chacun par deux des trois fonctions objectives. Sur chacune des figures, les solutions de la frontière Pareto du sous-problème sont marquées par un symbole carré et le reste des solutions

sont mises en évidence par le symbole '+'. Sur la Figure 6.2, la courbe de compromis dans le plan du coût total et du temps de séjour de l'eau peut être observée. Considérant le compromis entre les deux objectifs, on choisit une solution au point correspondant au rendement décroissant sur la courbe de compromis, car après ce point, il est y'a peu d'amélioration sur le temps de séjour de l'eau avec l'augmentation du coût total. De cette manière, la solution 1 est dans un premier temps identifiée à des fins de comparaison. Avec un raisonnement similaire, la solution 2 est identifiée au point correspondant au rendement décroissant sur la courbe de compromis entre le coût total et l'indice de résilience de la Figure 6.3. Enfin, la solution 3 représente un compromis entre l'indice de résilience et le temps de séjour de l'eau (Figure 6.4) et donc est sélectionné pour la comparaison.

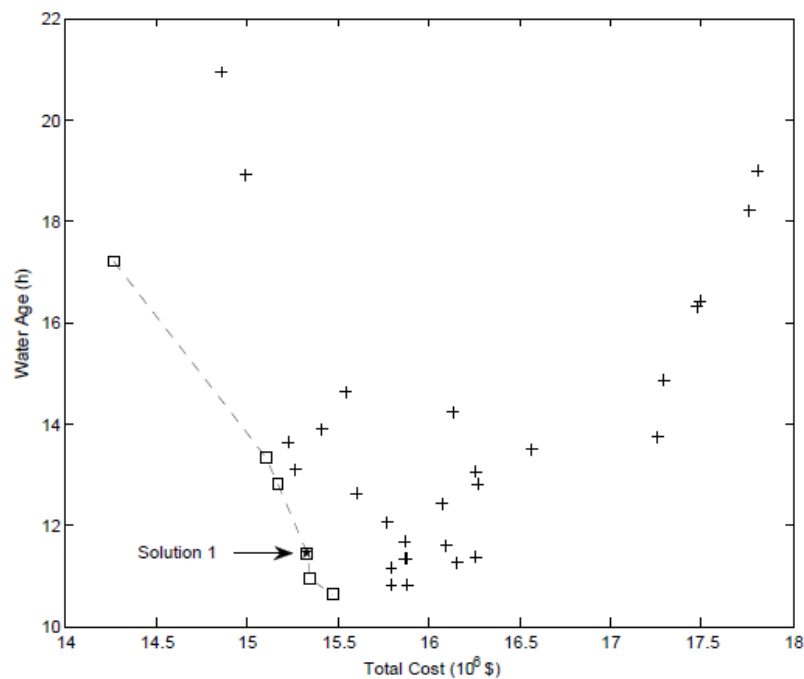


Figure 6.2 : Frontière de Pareto pour le coût total versus le temps de séjour de l'eau dans l'Alternative 1

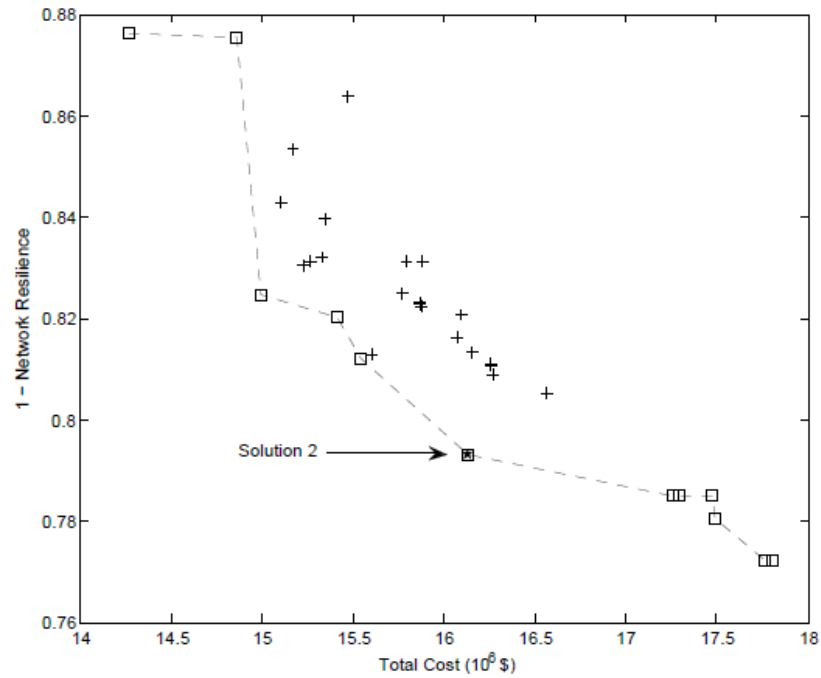


Figure 6.3 : Frontière de Pareto pour le coût total versus l'indice de résilience dans l'Alternative

1

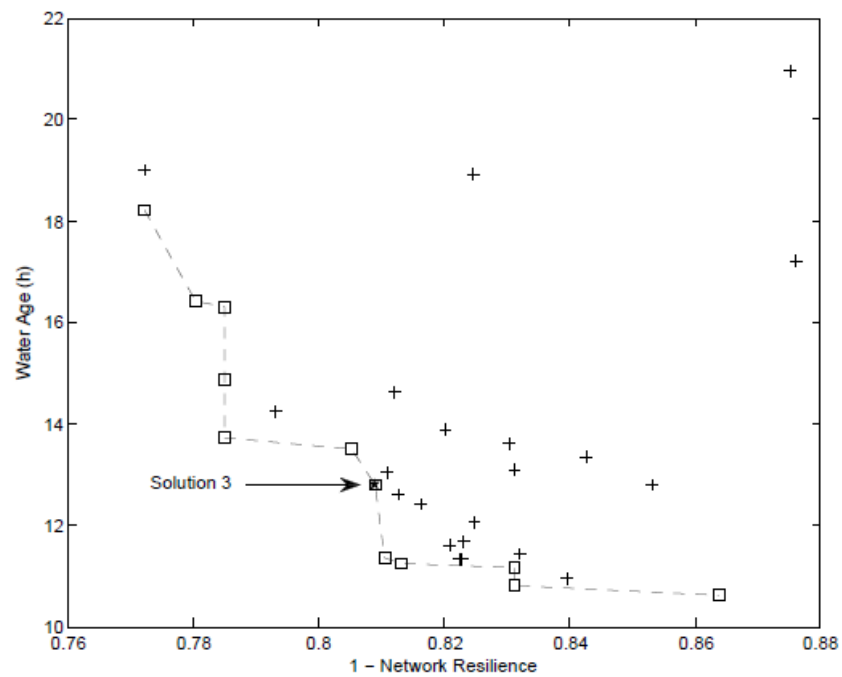


Figure 6.4 : Frontière de Pareto pour l'indice de résilience versus le temps de séjour de l'eau dans l'Alternative 1

Une inspection détaillée des valeurs des différentes fonctions objectives pour les trois solutions (voir Tableau 6.2) montre que la solution 2 offre un bon compromis en termes de coût total et cette solution est la meilleure du point de vue de l'indice de résilience du réseau. Le temps de séjour de l'eau pour chaque solution demeure faible. Avec ces considérations, la Solution 2 est recommandée dans cette alternative.

Tableau 6.2: Valeurs des fonctions objectives pour les six solutions sélectionnées

Numéro de la Solution	1	2	3	4	5	6	Lansey & Mays
Coût Total (10^6 \$)	15,33	16,13	16,27	12,30	12,62	13,46	13,33
Temps de séjour de l'eau (h)	11,44	14,23	12,80	6,01	22,04	7,58	38,87
Indice de résilience	0,17	0,21	0,19	0,10	0,13	0,16	0,08

Alternative 2

Dans cette alternative, le réservoir et par conséquent la conduite 4 sont supprimés du réseau. Le débit requis pour le réseau est pompé directement de la source et la pompe est alors active à temps plein. Seuls les diamètres des conduites sont alors variables, ce qui donne un nombre total de 16 variables pour le problème d'optimisation. La modélisation de ces variables est identique à celle adoptée dans l'alternative 1. L'algorithme génétique a été exécuté avec une taille de population de 50 solutions et un nombre de générations égal à 500. La frontière de Pareto identifiée par l'algorithme génétique à la fin des itérations contenait 40 solutions.

Comme dans les Figure 6.2 à 6.4, les Figures 6.5 à 6.7 montrent la projection de la frontière Pareto dans les trois plans. De la même manière que dans l'alternative 1, les Solutions 4, 5 et 6 sont respectivement identifiées sur les trois frontières de Pareto des Figures 6.5, 6.6 et 6.7 comme présentant de bons compromis.

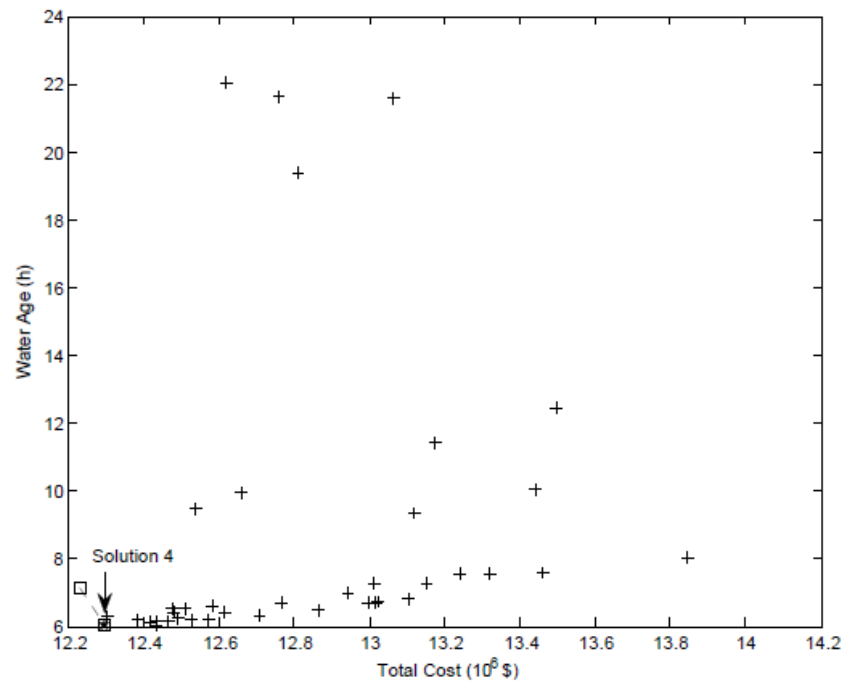


Figure 6.5 : Frontière de Pareto pour le coût total versus le temps de séjour de l'eau dans l'Alternative 2

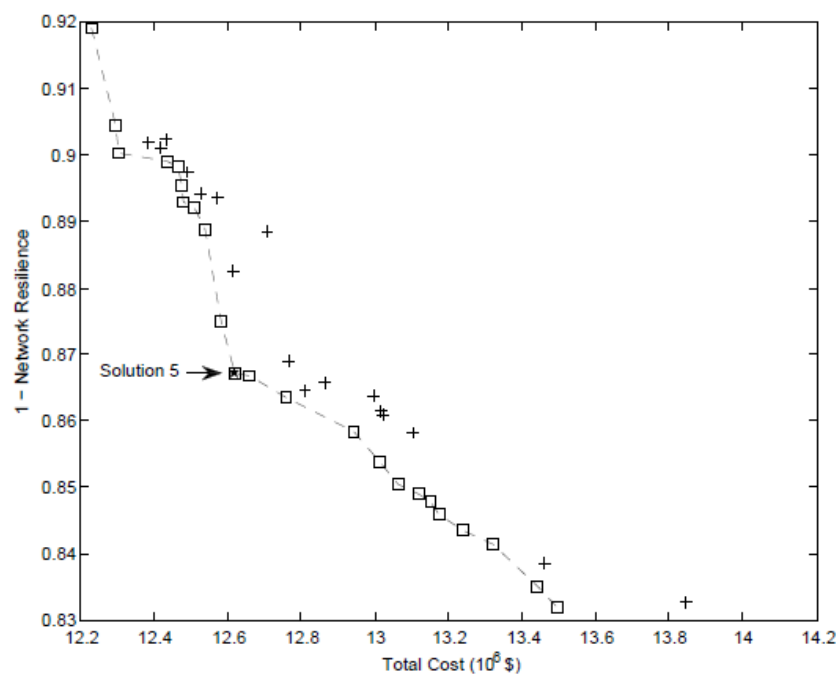


Figure 6.6 : Frontière de Pareto pour le coût total versus l'indice de résilience dans l'Alternative

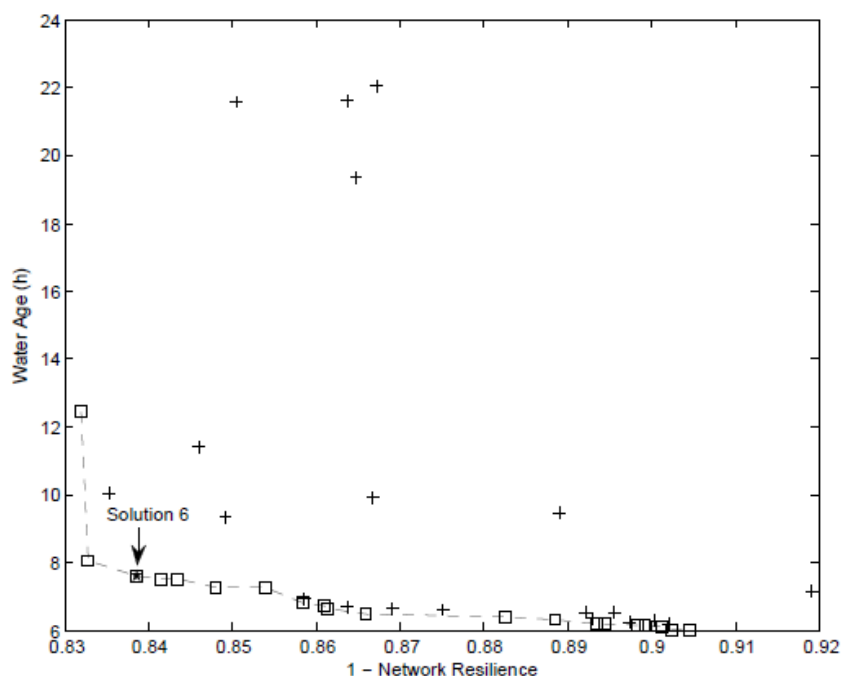


Figure 6.7 : Frontière de Pareto pour l'indice de résilience versus le temps de séjour de l'eau dans l'Alternative 2

Une inspection détaillée des valeurs des différentes fonctions objectives pour les trois solutions (voir Tableau 6.2) montre que la solution 6 est la meilleure en termes de l'indice de résilience du réseau et offre un bon compromis quant au temps de séjour de l'eau. Même si cette solution demeure relativement la plus chère, elle offre un bon compromis et est donc recommandée dans cette deuxième alternative.

Comparaison et alternative recommandée

Une analyse détaillée des deux alternatives révèle des solutions moins coûteuses dans la seconde alternative, mais moins performants en matière de résilience du réseau. En outre, le fait que la pompe fonctionne à temps plein dans la seconde alternative provoquerait une déconnection du réseau en cas de rupture de la station de pompage. Un réservoir, si présent, alimenterait le système dans un tel cas.

La Figure 6.8 montre les configurations de réseau pour les solutions recommandées 2 et 6 aussi bien que celle de Lansey et Mays (1989) pour la comparaison. Les périodes d'arrêt de la pompe dans la solution 2 sont : 3 ; 5 ; 15 et 18. Il est à noter que les périodes de temps vont de 1 à 24, chacune correspondant à l'heure de la journée type. La solution identifiée par Lansey et Mays

(1989) est converti conformément aux diamètres commerciaux et les équations de coût sont appliquées. De l'analyse du Tableau 6.2 et de la Figure 6.8, l'optimisation simultanée des trois fonctions objectives à savoir le coût total, l'indice de résilience du réseau et le temps de séjour de l'eau ont l'avantage d'offrir une meilleure qualité de l'eau et une bonne fiabilité en comparaison avec celui de l'optimisation d'une seule fonction objective qui est le coût total. La première solution avec un réservoir est plus fiable. Toutefois, le choix final dépend des priorités du décideur.

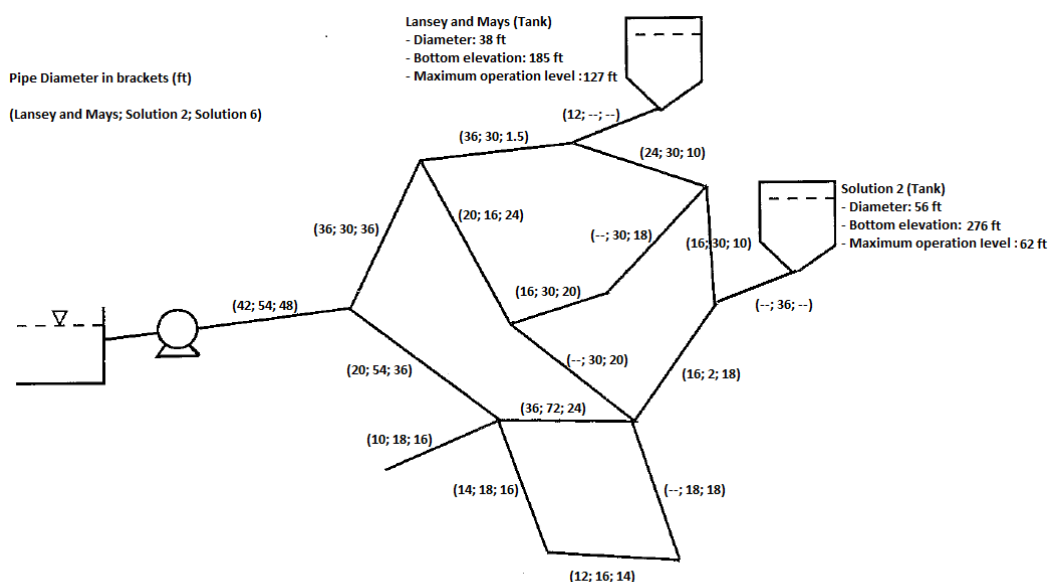


Figure 6.8 : Comparaison des résultats de conception: Lansey & Mays; Solution 2 et Solution 6

6.3.2 Approche stochastique sans corrélations

Dans cette approche de résolution du problème, les demandes de base aux nœuds sont supposées être des variables aléatoires sans aucune corrélation entre elles. La courbe de modulation journalière (fournie dans la description du problème) quant à elle reste invariable dans la journée type. L'échantillonnage « Latin Hypercube » est utilisée pour générer les demandes stochastiques qui par hypothèse suivent une distribution beta symétrique.

Les variables utilisées sont les mêmes que celles de l'Alternative 1 de l'optimisation déterministe précédente. L'indice de résilience est remplacé par le paramètre de robustesse (qui traduit la probabilité de dépassement des pressions minimales aux nœuds). La pénalité est

appliquée sur les valeurs moyennes de sortie. Une première étape d'optimisation déterministe est faite pour quelques générations et la population issue de la dernière génération est utilisée comme population initiale dans l'étape stochastique. Ceci permet de pouvoir commencer l'étape stochastique avec une population initiale moins aléatoire.

Tableau 6.3 : Résumé des résultats pour le réseau de l'exemple

Pipe number	Lansey & Mays	Proposed Model Solutions									
	(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	42	54	54	54	54	54	54	54	54	54	54
2	36	30	30	30	20	20	30	30	30	30	30
3	36	1,5	1,5	4	1,5	1,5	1,5	1,5	1	4	1,5
4	12	64	24	24	24	24	64	64	24	24	24
5	24	16	16	16	18	18	16	16	16	8	48
6	16	18	24	18	18	18	18	20	18	18	18
7	16	8	8	8	8	8	8	8	8	8	8
8	-	20	20	20	18	20	20	20	20	20	20
9	12	16	14	16	16	16	16	16	14	12	16
10	14	1,5	1,5	1,5	36	36	1,5	1,5	1,5	1,5	1,5
11	36	72	48	72	30	30	72	72	72	36	30
12	20	30	30	30	30	30	30	30	30	30	30
13	20	10	10	10	8	8	10	10	10	10	8
14	-	30	36	30	36	36	30	30	30	30	42
15	16	30	30	30	36	36	36	30	36	30	30
16	-	48	42	42	42	42	42	48	42	42	48
17	10	18	18	18	30	36	18	18	18	18	18
Pump off periods	-	3	3	3	3	3	3	3	3	3	3
	-	5	5	5	5	5	5	5	5	6	5
	-	10	6	6	6	6	6	10	6	10	6
	-	13	10	10	10	10	10	13	10	13	10
	-	15	13	13	13	13	13	15	13	15	13
	-	20	15	15	15	15	15	20	15	20	15
	-	-	20	20	20	20	20	-	20	-	20
	3	10	10	10	10	10	10	10	10	10	10
	38,26	73,0	73,0	73,0	73,1	73,1	73,1	73,0	73,0	73,0	73,0
	127,3	81,4	81,4	81,4	81,5	81,5	81,5	81,4	81,4	81,4	81,4
Tank location	3	10	10	10	10	10	10	10	10	10	10
Tank diameter (ft)	38,26	73,0	73,0	73,0	73,1	73,1	73,1	73,0	73,0	73,0	73,0
Tank maximum operation level (ft)	127,3	81,4	81,4	81,4	81,5	81,5	81,5	81,4	81,4	81,4	81,4
Tank bottom elevation (ft)	185	218	218	218	218	218	218	218	218	218	218
Total cost (10 ⁶ \$)	13,33	16,34	15,80	15,87	15,99	16,14	16,15	16,37	15,91	15,40	16,59
Water age (h)	43,33	24,29	15,61	14,93	14,62	15,53	19,02	24,14	14,88	16,82	15,18
Robustness parameter	1,03	0,608	1,183	0,809	1,116	1,124	0,825	0,585	1,08	0,92	0,551
Robustness (%)	85	72	88	79	86	87	79	72	86	82	71

Afin de déterminer les caractéristiques du compromis entre le coût total, le paramètre de robustesse et le temps de séjour de l'eau, l'algorithme génétique a été exécuté avec une taille de

population de 50 solutions et un nombre de générations égal à 2000 dans l'étape déterministe et 200 générations dans l'étape stochastique. Le nombre d'échantillons utilisés pour la simulation stochastique est 30. La fiabilité des résultats malgré ce nombre faible vient de la technique « Latin Hypercube » utilisée.

La frontière de Pareto identifiée par l'algorithme génétique à la fin des itérations contenait 47 solutions. A partir de ces solutions, 10 solutions potentielles ont été extraites et présentées au Tableau 6.3, où chaque colonne de (1) à (10) représente une solution au problème.

Rappelons que les périodes de temps vont de 1 à 24, chacune correspondant à l'heure de la journée.

Une analyse détaillée des solutions identifiées et reportées dans le Tableau 6.3 montre un certain nombre de points intéressants. En effet, toutes les solutions identifiées sont associées à une très bonne qualité de l'eau (temps de séjour entre 14h et 24h) tandis que la solution de Lansey et Mays (1989) est associée à un temps de séjour de plus de 43h. La robustesse du réseau va de 71 à 88%. Cela signifie que les solutions reportées dans le Tableau 6.3 satisfont à l'exigence de pression minimum à tous les nœuds avec une probabilité supérieure à 70%. En termes de conception du réservoir, toutes les solutions identifiées sont similaires : elles suggèrent de placer le réservoir sur le nœud 10 avec une élévation du radier de 218 pieds, un diamètre de 73 pieds et un niveau d'opération maximale de 81 pieds. D'un point de vue technique, cinq des dix solutions ne seraient pas acceptables, car elles suggèrent des diamètres de conduite allant jusqu'à 72 pouces. Ces solutions inacceptables sont (1), (3), (6), (7) et (8). Les cinq solutions restantes proposent un diamètre maximum de 54 pouces, ce qui est techniquement valable. Pour éviter des solutions de l'algorithme génétique avec des diamètres excessivement élevés, il est nécessaire d'imposer une limite supérieure sur les possibilités de choix de diamètres des conduites lors de la création du génome. Les solutions nécessitant une analyse plus approfondie seraient alors être (2), (4) ou (5). Cependant, ces solutions sont un peu plus chères que celle proposée par Lansey et Mays, mais sont plus robustes et plus fiables en termes de qualité de l'eau. Étant donné que la solution (5) est différente de la solution (4) uniquement par le diamètre de la conduite 8, la Figure 6.9 compare les résultats obtenus pour les solutions (0), (2) et (4). Le triplet dans chaque parenthèse à côté de chaque conduite correspond aux diamètres proposés pour chacune des trois solutions respectivement. Il paraît évident que les solutions générées par l'optimisation

stochastique des trois fonctions objectives à savoir le coût total, la robustesse et le temps de séjour de l'eau présentent l'avantage d'une meilleure qualité de l'eau et de fiabilité par rapport à celle de l'optimisation mono-objective du coût total.

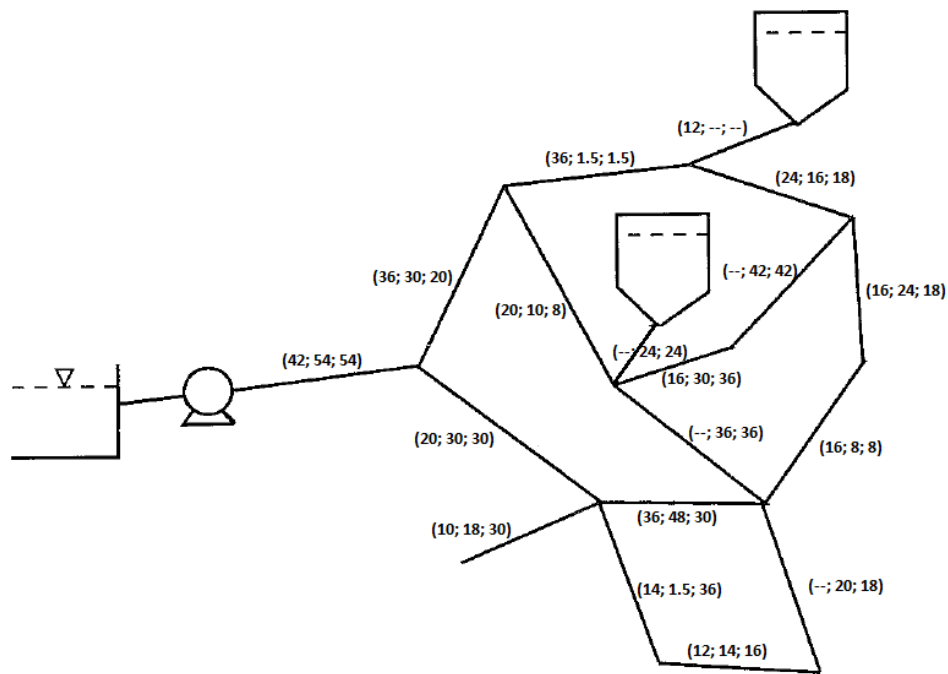


Figure 6.9 : Comparaison des résultats pour les solutions (0), (2) et (4)

6.3.3 Approche stochastique avec corrélations

Dans cette approche de résolution du problème, les demandes de base aux nœuds sont supposées être des variables aléatoires corrélées entre elles. Il est aussi décidé ici de réduire le nombre de variables le plus possible afin d'en évaluer également l'impact. Pour ce faire, la démarche décrite dans l'article 2 (Équations (3.5) et (3.6)) est utilisée pour évaluer la réserve d'équilibre du réservoir. La demande totale horaire déterministe du réseau est alors évaluée pour la journée type. Les calculs détaillés sur chaque pas de temps horaire sont reportés au Tableau 6.4.

Tableau 6.4 : Calcul de la réserve d'équilibre et du débit de pompage pour le réseau de l'exemple

Période	Demande (pi ³ /s)	Demande (pi ³)	Demande Cumulée (pi ³)	Pompage (pi ³)	Stockage (pi ³)
					0
1	12,05	86 736	86 736	131 074	44 339
2	12,05	86 736	173 471	131 074	88 677
3	12,05	86 736	260 207	131 074	133 016
4	12,05	86 736	346 942	131 074	177 355
5	12,05	86 736	433 678	131 074	221 693
6	12,05	86 736	520 413	131 074	266 032
7	21,58	155 347	675 760	131 074	241 759
8	21,58	155 347	831 108	131 074	217 486
9	21,58	155 347	986 455	131 074	193 213
10	21,58	155 347	1 141 802	131 074	168 940
11	21,58	155 347	1 297 149	131 074	144 667
12	21,58	155 347	1 452 496	131 074	120 394
13	14,02	100 976	1 553 472	131 074	150 493
14	14,02	100 976	1 654 448	131 074	180 591
15	14,02	100 976	1 755 423	131 074	210 690
16	14,02	100 976	1 856 399	131 074	240 788
17	14,02	100 976	1 957 375	131 074	270 887
18	14,02	100 976	2 058 350	131 074	300 985
19	25,17	181 238	2 239 589	131 074	250 821
20	25,17	181 238	2 420 827	131 074	200 657
21	25,17	181 238	2 602 066	131 074	150 493
22	25,17	181 238	2 783 304	131 074	100 328
23	25,17	181 238	2 964 542	131 074	50 164
24	25,17	181 238	3 145 781	131 074	0
		Q _{moy} (pi ³ /h)	131 074	Volume du Réservoir (pi ³)	300 985
		Q _{moy} (pi ³ /s)	36,41		

Cette démarche a permis d'obtenir un volume d'équilibre d'environ $300\,000\text{ pi}^3$ pour le réservoir pour un débit de pompage constant de $37\text{ pi}^3/\text{s}$. Un ratio diamètre/hauteur de 0,8 est adopté pour le réservoir, incluant une revanche de 10% du volume total. Cette hypothèse permet de fixer le diamètre du réservoir à 70 pieds conduisant à un niveau maximum d'opération de 80 pieds. Le débit de la pompe est alors fixé à $37\text{ pi}^3/\text{s}$.

Pour une pression maximale de 500 kPa, l'application de l'approche l'article 2 (Équations (3.3) et (3.4)) indique que l'altitude du radier du réservoir devrait être entre 215 pieds et 300 pieds. Une altitude moyenne du radier de 260 pieds est alors fixée pour le réservoir. Lorsque le réservoir est plein, l'altitude du niveau d'eau est alors de 340 pieds. Le niveau de la source (point d'aspiration de la pompe) étant à 50 pieds, pour pouvoir pomper l'eau de la source et remplir le réservoir, il faut une charge de 290 pieds (si les pertes de charge sont négligées). La hauteur manométrique de la pompe est alors fixée à 290 pieds. Au lieu d'adopter une politique d'opération de la pompe sur une base horaire comme dans les approches précédentes, il est décidé ici de baser l'opération sur le niveau du réservoir : la pompe démarre quand le niveau d'eau dans le réservoir descend en bas d'un seuil (5 pieds fixés dans cet exemple) et s'arrête quand le réservoir est plein (80 pieds).

Avec le traitement préliminaire précédant, les seules variables considérées dans cette approche sont les diamètres des conduites et l'emplacement du réservoir. Pour permettre le choix de diamètres raisonnables par l'algorithme génétique, la plage de sélection des diamètres commerciaux a été limitée à des valeurs allant de 6 pouces à 48 pouces (13 alternatives conformément au Tableau 6.1). Avec ces 13 alternatives, le diamètre est modélisé comme un entier prenant des valeurs entre 10 et 22. Le nombre de variables est donc : 17 pour les conduites et 1 pour le réservoir ; ce qui donne un total de 18 variables pour le problème d'optimisation.

Pour les fonctions objectives, le temps de séjour de l'eau est remplacé par le paramètre W_{Anet} défini dans l'article 3 (Équation (4.11)) avec un seuil $W_{A_{th}}$ de 24h, ce qui signifie que le temps de séjour d'un nœud donné entre dans les calculs seulement si ce temps est supérieur à 24h. Le coût total et le paramètre de robustesse restent ceux utilisés dans l'approche précédente. Les mêmes contraintes et paramètres d'échantillonnage que ceux de l'approche stochastique sans corrélation sont utilisés ici. Une corrélation de 0.5 a été fixée pour les demandes aux nœuds.

Tableau 6.5 : Résumé des résultats stochastiques avec corrélation pour le réseau de l'exemple

Pipe number	Proposed Model Solutions						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	24	24	24	24	24	24	24
2	36	36	36	36	36	36	36
3	14	8	6	8	12	6	12
4	24	24	24	24	24	24	24
5	10	10	10	10	10	10	10
6	12	12	12	12	12	12	12
7	48	48	24	48	48	24	42
8	20	20	8	20	8	8	8
9	18	18	18	18	18	18	18
10	12	12	36	12	14	36	12
11	14	14	20	14	36	20	14
12	18	18	48	18	48	48	48
13	42	42	42	42	42	42	48
14	12	12	12	12	12	12	12
15	42	48	42	42	42	42	42
16	42	42	30	30	30	30	42
17	24	48	12	12	48	30	24
Tank location	2	2	2	2	2	2	2
Total cost (10^6 \$)	17,85	18,31	17,80	17,34	18,99	18,08	18,52
WA _{net} (h)	0,248	0,355	0,013	0,197	0,362	0,031	0,221
Robustness parameter	0,983	0,987	1,122	0,940	1,238	1,122	1,191
Robustness (%)	84	84	87	83	89	89	88

Il est à noter que le volume du réservoir étant fixé de façon déterministe, ce sont les diamètres des conduites qui devront tenir compte de l'aspect stochastique du problème. L'algorithme génétique a été exécuté avec une taille de population de 30 solutions et un nombre de générations 100 générations. L'optimisation commence directement de façon stochastique et aucune étape préliminaire déterministe n'est effectuée ici comme ce fut le cas précédemment.

La frontière de Pareto identifiée par l'algorithme génétique à la fin des itérations contenait 8 solutions dont une répétée. Les 7 solutions non répétées sont présentées au Tableau 6.5, où chaque colonne de (1) à (7) représente une solution au problème.

Rappelons que les périodes de temps vont de 1 à 24, chacune correspondant à l'heure de la journée.

Une analyse détaillée des solutions identifiées et reportées dans le Tableau 6.5 montre que l'intégration des corrélations dans les demandes aux nœuds conduit à des solutions plus chères mais aussi plus fiables. La robustesse du réseau va de 83 à 89%. Cela signifie que les solutions reportées dans le Tableau 6.5 satisfont à l'exigence de pression minimum à tous les nœuds avec une probabilité supérieure à 80%. Le paramètre traduisant le temps de séjour va de 0,013 (Solution (3)) à 0,355 (Solution (2)), avec une moyenne de 0,204, ce qui reste faible. En termes d'emplacement du réservoir, toutes les solutions suggèrent de placer le réservoir sur le nœud 2 (l'un des nœuds les plus élevés du réseau). La solution (4) est recommandée dans cette approche pour son avantage d'être le moins cher et présentant un compromis quant à la qualité de l'eau. Son paramètre de robustesse (83%), bien qu'étant la plus faible parmi les solutions trouvées demeure une valeur largement acceptable. La réduction du nombre de variables et la limitation des alternatives de diamètre de conduites ont permis d'arriver très rapidement à de bonnes solutions.

6.4 Conclusion

L'application des différentes approches proposées sur un réseau pris dans la littérature permet de tirer plusieurs conclusions. Sur la base des solutions générées, le classement des approches par coût croissant se présente comme suit : (1) multiobjective déterministe sans réservoir; (2) mono-objective déterministe avec réservoir; (3) multiobjective déterministe avec réservoir; (4) multiobjective stochastique avec réservoir sans corrélations et (5) multiobjective stochastique

avec réservoir et corrélations. On peut alors dire d'une part que la solution avec réservoir est moins chère que celle sans réservoir. D'autre part, l'approche déterministe est moins coûteuse que celle stochastique. Au sein de l'approche stochastique elle-même, cela revient plus cher si on choisit d'inclure les corrélations dans les demandes aux nœuds. Mais en termes de fiabilité du réseau, le classement des cinq approches est inversé : les plus chères sont les plus fiables, ce qui est tout à fait logique et d'où la notion de compromis. Quelque soit l'approche utilisée, l'étape de formulation du problème (choix des variables et bornes) est très importante car le fardeau réservé à l'algorithme génétique pour produire des solutions acceptables devient important s'il n'y a pas un traitement préliminaire qui est fait pour réduire le nombre de variables. Enfin on peut dire que la démarche de résolution proposée dans cette thèse s'applique bien avec succès sur les projets de conception et de gestion des réseaux de distribution d'eau potable.

CHAPITRE 7 DISCUSSION GÉNÉRALE

La revue critique de la littérature (article 1) a montré que l'intégration des réservoirs dans la conception optimale des réseaux de distribution d'eau potable n'est pas un aspect nouveau en soi. Ce qui fait la différence est la manière dont les réservoirs sont intégrés dans le processus d'optimisation notamment les variables de décision associées à ces réservoirs ainsi que les hypothèses posées pour résoudre le problème. L'un des aspects innovants du projet est le fait qu'une emphase est mise sur la conception des réservoirs comme partie intégrante du processus d'optimisation, vu l'impact significatif que ces réservoirs ont sur les performances globales des réseaux de distribution. L'optimisation des différentes composantes du réseau est faite simultanément vu l'interdépendance entre elles.

En termes de qualité de l'eau, la motivation pour son intégration dans le processus d'optimisation est d'ordre préventif : il est en effet plus judicieux de prendre en compte la qualité de l'eau dès la conception, que de résoudre les problèmes une fois les infrastructures réalisées. Plusieurs paramètres peuvent être utilisés pour caractériser la qualité de l'eau : la croissance de microorganismes, la décroissance du chlore résiduel et le temps de résidence hydraulique. En modélisation numérique, la plupart des modèles de qualité d'eau utilisent le temps de résidence hydraulique pour calculer les concentrations de différents composés. L'hypothèse faite dans cette étude est que plus le temps de séjour est court, meilleure est la qualité de l'eau. C'est ce qui justifie l'utilisation de cet indicateur comme un des objectifs à minimiser lors de la conception optimale des réseaux de distribution. Bien qu'un temps de séjour faible ne soit pas synonyme d'une bonne qualité de l'eau, minimiser ce temps permet de réduire la détérioration de qualité initiale de l'eau. Un autre aspect innovant du projet est la prise en compte des aspects de la qualité de l'eau dans l'optimisation en environnement stochastique des réseaux de distribution d'eau potable. Cette inclusion permet d'étendre la conception stochastique aux réseaux contenant un ou plusieurs réservoirs vu que ces réservoirs introduisent inévitablement des problèmes de qualité de l'eau dans les réseaux.

Pour ce qui est de l'approche de formulation du problème, il a été réalisé que l'approche de formulation multiobjective est la plus adaptée, contrairement à la formulation mono-objective. Cette formulation conduit à des réseaux plus performants moyennant un compromis à faire quant au coût résultant. Parmi les multiples avantages de cette approche multiobjective, il convient de

mentionner que la proposition d'un ensemble de solutions alternatives répond mieux aux besoins des concepteurs qui peuvent comparer les solutions et opérer un choix en fonction de leurs priorités. Ceci explique l'orientation du projet vers une approche multiobjective.

Concernant la méthode de résolution à adopter, les algorithmes génétiques se sont révélés efficaces pour traiter le problème de conception optimale des réseaux de distribution d'eau potable. Un des avantages de ces algorithmes est leur capacité à traiter adéquatement les variables discrètes qui sont largement sollicitées dans la conception des réseaux. Aussi les résultats obtenus par ces algorithmes sont moins sensibles aux paramètres spécifiques utilisés dans leur implémentation. Par contre, ils nécessitent beaucoup d'évaluations des fonctions objectives avant de donner de bons résultats, mais ils permettent de traiter des réseaux complexes. Dans le cadre du projet de recherche, le code est implémenté pour être utilisé avec un algorithme génétique.

Par rapport à l'approche à utiliser sur les données de simulation, chacune des deux approches à savoir l'approche déterministe et l'approche stochastique avait des avantages et des inconvénients. Que ce soit dans l'une ou dans l'autre approche, il était possible de prendre en compte l'incertitude liée aux demandes en eau des nœuds. Cette prise en compte avait pour but de réduire l'écart observé entre la pratique sur le terrain et les modèles mathématiques d'optimisation proposés dans la littérature. L'approche déterministe est plus rapide en termes de simulations et permet donc d'obtenir des solutions acceptables en des temps raisonnables. Cet avantage rend l'approche mieux applicable sur de grands réseaux complexes. Par contre, bien que des marges de manœuvre soient prévues dans cette approche pour pouvoir répondre à des cas non prévus lors de la conception, les données spécifiques utilisées pour la simulation influencent les résultats obtenus. L'un des avantages majeurs de l'approche stochastique est qu'elle permet de tester le réseau dans de larges éventails de combinaisons de demandes de pointes. Ceci la rend plus réaliste par rapport à l'approche déterministe. L'approche stochastique fournit également au décideur l'information sur la performance statistique du réseau (probabilité d'échec par exemple). Par contre elle coûte très chère en termes de temps de calcul et se voit donc limitée sur la grandeur des réseaux pouvant être couverts.

Concernant les résultats, il est maintenant question d'étudier si les solutions proposées sont réellement implantables sur le terrain. L'objectif du projet est le développement d'un modèle

d'optimisation multiobjective des réseaux de distribution d'eau potable capable de prédire la nécessité ou non d'introduction d'un ou de plusieurs réservoirs dans le réseau ainsi que l'emplacement de ces éventuels réservoirs. Les approches proposées (articles 2 et 3) ont été appliquées à trois études de cas :

- Étude de cas 1 : l'extension et la réhabilitation du réseau de la municipalité de D-Town en utilisant l'approche déterministe (article 2).
- Étude de cas 2 : la réhabilitation du réseau de la municipalité de Saint-Jérôme en utilisant un dimensionnement en environnement stochastique (article 3).
- Étude de cas 3 : le dimensionnement d'un réseau à quatre boucles représentatif des réseaux de petites villes. Deux configurations pour ce réseau ont été étudiées : la première inclut un réservoir et la seconde exclut tout réservoir. Les deux approches (déterministe et stochastique) ont été appliquées et au sein de l'approche stochastique, les options d'intégration ou non de la corrélation entre les demandes aux nœuds ont été également étudiées.

Les résultats proposés sont fournis à partir du modèle développé couplé avec le modèle de simulation numérique EPANET. L'étude de cas 1 a montré que l'approche proposée permet l'ajout d'un réservoir dans un projet d'extension de réseau sans nécessairement recourir à l'étape d'optimisation pour déterminer sa capacité. Ceci permet d'alléger les variables d'optimisation pour permettre à l'algorithme d'évoluer rapidement vers des solutions praticables sur le terrain. L'étude de cas 2 a montré que le modèle développé permet une conception efficace de la capacité et de l'emplacement des réservoirs tout en améliorant la qualité de l'eau dans le réservoir optimisé. L'étude de cas 3 quant à elle a montré que des deux alternatives possibles analysées, celle excluant tout réservoir était la moins chère mais également la moins fiable en termes de résilience du réseau. La nécessité dans ce cas d'inclure un réservoir dépend alors des contraintes supplémentaires du décideur notamment l'aspect financier. Les solutions proposées en environnement stochastique sont plus chères que celles obtenues avec l'approche déterministe. Et même au sein de l'approche stochastique, lorsque les corrélations sont considérées entre les demandes aux nœuds, les solutions résultantes sont plus chères que lorsque ces demandes sont traitées comme indépendantes. La remarque générale qui en découle est donc que plus on cherche à s'approcher des conditions réalistes lors de la simulation, plus le coût du

réseau est élevé. Ce qui est normal vu le caractère aléatoire des données utilisées pour cette simulation. Il faut cependant noter que l'approche stochastique est difficile à implémenter vu l'ensemble des données qu'elle nécessite de collecter sur le terrain. En effet, la matrice de corrélations à elle seule nécessite un historique des consommations simultanées de tous les consommateurs individuels sur une longue période de temps ainsi que la variation dans le temps de ces consommations. Ces données ne sont pas généralement disponibles.

Le Tableau 7.1 donne une synthèse des résultats significatifs obtenus dans la thèse en relation avec les approches et paramètres utilisés tels que présentés au Tableau 1.1.

Tableau 7.1 : Résultats obtenus dans la thèse en relation avec les approches et paramètres utilisés

Approche ou paramètre	Résultat
Multiobjective	A permis d'analyser en moyenne environ six solutions sur la frontière Pareto pour chaque approche de conception utilisée en vue d'un compromis; a conduit à une réduction du temps de séjour d'environ 70% moyennant une augmentation de coût de 15% sur le cas du réseau à quatre boucles
Temps de séjour	Son intégration en environnement stochastique a permis d'étendre la conception stochastique aux réseaux contenant un ou plusieurs réservoirs; a conduit à une réduction du temps de séjour d'environ 70% sur le cas du réseau de Saint-Jérôme
Déterministe	A été 40 fois plus rapide et 2% moins coûteux que l'approche stochastique sur le cas du réseau à quatre boucles; une simulation stochastique sur les résultats déterministes a donné un échec sur un pourcentage d'échantillons allant jusqu'à 50
Stochastique	A été 40 fois plus lent et 2% plus coûteux que l'approche déterministe mais a permis de trouver des solutions avec une probabilité d'échec inférieure à 30% sur le cas du réseau à quatre boucles
Corrélations	Son intégration dans les demandes aux noeuds a conduit en moyenne à une conception 7% plus fiable mais 13% plus chère sur le cas du réseau à quatre boucles
Étude de cas 1: D-Town	Modèle testé sur un projet mixte: extension et réhabilitation; a permis de couvrir l'optimisation des conduites et l'opération des vannes tout en réduisant le nombre de variables d'environ 90 %
Étude de cas 2: Saint-Jérôme	Modèle testé sur un projet de réhabilitation; a permis de couvrir l'optimisation du réservoir et la capacité de la pompe
Étude de cas 3: Réseau à quatre boucles	Modèle testé sur la conception d'un nouveau réseau; a permis de couvrir l'optimisation des conduites, du réservoir, et l'opération de la pompe; a permis de comparer les différentes approches de conception des réseaux

CHAPITRE 8 CONCLUSIONS ET RECOMMANDATIONS

8.1 Conclusions générales

Le but de l'étude consistait à mettre en place une méthodologie efficace permettant d'intégrer les réservoirs dans l'optimisation multiobjective des réseaux de distribution d'eau potable avec comme objectifs de maximiser la qualité de l'eau et divers mesures de fiabilité tout en minimisant les coûts d'installation et d'exploitation des réseaux conçus. Le problème étudié était de trouver la conception et la politique d'opération optimale d'un réseau (avec ou sans réservoir) qui est capable de fournir les demandes en eau prévues tout en respectant les exigences quantitatives, qualitatives et hydrauliques. Les travaux réalisés ont permis de dégager les conclusions suivantes :

- L'étude des approches existantes pour la conception des réservoirs a été une étape déterminante quant à l'évaluation des besoins. Elle a permis de constater l'existence de deux grandes familles d'approche : d'une part celles basées sur l'expérience en génie et les guides techniques et d'autres parts celles basées sur les modèles mathématiques d'optimisation. Un écart important a été observé entre ces deux approches dont l'une des causes peut être cherchée dans le manque de collaboration entre les deux parties et la difficulté pour les modèles mathématiques d'intégrer les vrais enjeux de terrain. La nature stochastique du problème lui-même explique aussi en partie les divergences observées.
- La méthodologie s'appuie sur un algorithme d'optimisation qui a été mis au point au cours de l'étude. Cet algorithme, au stade de logiciel pilote, a été testé sur trois études de cas et a prouvé son efficacité pour déterminer de bonnes solutions. L'approche multiobjective offre beaucoup plus de compromis entre les divers objectifs et aussi plus d'alternatives aux gestionnaires des RDEP. La prise en compte de la qualité de l'eau permet une intégration efficace des réservoirs dans le processus d'optimisation. Bien que n'étant pas la plus réaliste, l'approche déterministe propose des solutions qui peuvent être obtenues en beaucoup moins de temps comparé à l'approche stochastique qui peut s'avérer non applicable sur des réseaux très grand et complexes.

- Une analyse rigoureuse des résultats, à l'aide d'une méthode multicritères, permet de choisir une solution qui convient le mieux à chaque situation. Il est essentiel de bien prendre en compte à ce stade les attentes des gestionnaires en termes de qualité d'eau, de coûts d'infrastructures et d'opérations pour proposer une solution qui réponde parfaitement à leurs besoins.

8.2 Contributions

8.2.1 Articles publiés dans des revues avec comité de lecture

- Batchabani, E. and Fuamba, M. (2014b). "Optimal Tank Design in Water Distribution Networks: Review of Literature and Perspectives." *Journal of Water Resources Planning and Management*, 140(2), 136–145.
- *Batchabani, E. and Fuamba, M. (2013a). "Discussion of "Water Distribution System Analysis: Newton-Raphson Method Revisited" by M. Spiliotis and G. Tsakiris." *Journal of Hydraulic Engineering*, 139(8), 917–918.
- *Batchabani, E., Jalliffier-Verne, I., and Fuamba, M. (2012). "Discussion of "Theory-Based SCS-CN Method and Its Applications" by W. H. Chung, I. T. Wang, and R. Y. Wang." *Journal of Hydrologic Engineering*, 17(2), 354–355.

*N.B: Bien que l'article ne soit pas en lien avec le sujet de la thèse, il fait partie des contributions générales à l'avancement de la recherche.

8.2.2 Articles soumis et en cours de révision

- Batchabani, E. and Fuamba, M. (2013b). "Multi-objective design and decision-making process in water distribution systems." *Engineering Optimization*, Submitted.
- Batchabani, E. and Fuamba, M. (2014a). "Multi-objective stochastic design of water distribution systems including water quality." *Journal of water resources planning and management*, Submitted.

8.2.3 Principales contributions à la recherche et au développement

8.2.3.1 Orientation du choix d'une approche appropriée

En plus d'une analyse critique structurée des différentes approches incluant la considération des réservoirs dans l'optimisation des réseaux, un protocole est proposé afin de permettre l'identification de la meilleure approche ou du type de modèle approprié en fonction d'un certain nombre de critères prédéterminés. L'originalité de ce travail tient au fait que l'analyse se limite aux modèles incluant explicitement les réservoirs dans le processus d'optimisation.

8.2.3.2 Proposition de paramètres permettant de généraliser la conception

L'un des soucis majeurs du projet était de proposer un modèle dont le champ d'application soit large. C'est ce qui a conduit à la proposition de certains paramètres de références pouvant permettre la généralisation de l'approche de conception. Ces paramètres sont : (1) le nombre de zones de pression; (2) les élévations minimales et maximales des réservoirs (3) le diamètre de référence pour les conduites; (4) le débit et la hauteur manométrique de référence pour les pompes et (5) la sélection des paramètres de l'algorithme génétique. Cette approche est utile pour les applications pratiques de l'algorithme d'optimisation aux situations réelles. L'originalité de cette approche tient au fait que ces paramètres n'ont pas été utilisés par le passé comme des balises du processus d'optimisation.

8.2.3.3 Développement d'un modèle d'optimisation déterministe et stochastique

Un modèle numérique pour l'optimisation des réseaux de distribution d'eau potable est proposé dans cette étude afin d'aider les gestionnaires dans leurs projet d'installation ou de maintenance de réseaux de distribution. Une version de ce modèle est implémentée en environnement déterministe et l'autre en stochastique pour permettre une comparaison des deux approches de conception et en même temps pour donner plus de choix aux utilisateurs. La qualité de l'eau a été considérée comme faisant partie des objectifs de la conception optimale des réseaux dans les deux approches. L'originalité de ce travail tient au fait que la qualité de l'eau n'a pas été considérée par le passé comme objectif d'optimisation en environnement stochastique. Le modèle proposé n'est pas encore à son stade final facilement utilisable par les gestionnaires. Par conséquent il nécessite des améliorations quant à sa convivialité. Il permettra à terme

l'optimisation des éléments du réseau de distribution en explorant adéquatement l'ensemble des solutions pour proposer celles qui sont optimales. Il sera interactif et l'utilisateur aura le choix sur l'approche déterministe ou stochastique. L'utilisateur pourra également choisir le type d'opération qu'il souhaite pour les pompes ou imposer ses propres bornes quant à la plage des variables d'optimisation.

8.3 Recommandations

Quelques difficultés rencontrées au cours de la modélisation numérique sont relevées ci-après afin d'indiquer dans quelles directions les études futures devront s'orienter:

- *Temps de calcul* : plusieurs facteurs influencent ce temps de calcul : (1) l'algorithme génétique; (2) les fonctions d'évaluation et (3) la simulation EPANET. En général le temps de calcul est assez long pour la résolution de ce type de problèmes, vu que plusieurs milliers d'évaluations doivent être effectués. Parmi les trois facteurs mentionnés, il a été remarqué que la simulation EPANET était la plus longue en temps de calcul et particulièrement la simulation de la qualité de l'eau. En effet cette dernière, contrairement à la simulation hydraulique, requiert un pas de temps beaucoup plus petit et une durée de simulation beaucoup plus longue. Cette difficulté réduit la possibilité d'appliquer le modèle sur de grands réseaux. Il est donc recommandé d'explorer des alternatives pour sous-traiter l'évaluation même approximative de la qualité de l'eau. Une autre piste pouvant améliorer le temps de calcul serait d'étudier la possibilité de paralléliser le code pour répartir les tâches de calcul sur les différents cœurs de l'ordinateur si celui-ci en a plus d'un.
- *Convivialité* : le modèle est actuellement un programme sous forme de code C++ dont l'utilisation est un processus lourd et complexe. Le développement d'un programme complet, doté d'une interface élaborée permettrait de faciliter l'utilisation de ce dernier par les concepteurs de réseaux de distribution d'eau potable. Cette interface permettrait à l'étude de passer du stade de recherche à un stade d'utilisation courante.
- *Études de cas* : l'étude actuelle n'ayant pas fait l'objet de plusieurs tests grandeur nature, il est permis de penser que l'application de cette méthodologie à des projets pilotes

permettrait de recueillir des données pertinentes afin de valider le modèle de manière efficace.

- *Durée totale de simulation hydraulique* : Étant donné que dans le code actuel, la durée de simulation doit être personnalisée constamment, il est recommandé dans les futures investigations d'étudier la possibilité de généralisation de cette durée pour être capable de traiter des problèmes qui exigeraient des durées quelconques sans besoin d'intervention supplémentaire dans le code.

RÉFÉRENCES

- Afshar, M. H., Akbari, M., & Marino, M. A. (2005). Simultaneous layout and size optimization of water distribution networks: engineering approach. *Journal of Infrastructure Systems*, 11(4), 221-230.
- Ahn, C. W. (2006). *Advances in evolutionary algorithms: theory, design and practice* (Vol. 18): Springer Verlag.
- Alperovits, E., & Shamir, U. (1977). Design of optimal water distribution systems. *Water Resources Research*, 13(6), 885-900.
- Aly, A. H., & Wanakule, N. (2004). Short-term forecasting for urban water consumption. *Journal of water resources planning and management*, 130(5), 405-410.
- AWWA. (2008). *Distribution System Requirements for Fire Protection - Manual of Water Supply Practices, M31 (4th Edition)*. Denver, Colorado: American Water Works Association.
- Babayan, A., Kapelan, Z., Savic, D., & Walters, G. (2005). Least-cost design of water distribution networks under demand uncertainty. *Journal of water resources planning and management*, 131(5), 375-382.
- Babayan, A., Savic, D., Walters, G., & Kapelan, Z. (2007). Robust Least-Cost Design of Water Distribution Networks Using Redundancy and Integration-Based Methodologies. *Journal of water resources planning and management*, 133(1), 67-77.
- Basile, N. (2008). *Optimization of the design, location and management of water tank in water distribution systems*. Ecole Polytechnique de Montréal.
- Batchabani, E., & Fuamba, M. (2013a). Discussion of "Water Distribution System Analysis: Newton-Raphson Method Revisited" by M. Spiliotis and G. Tsakiris. *Journal of Hydraulic Engineering*, 139(8), 917-918.
- Batchabani, E., & Fuamba, M. (2013b). Multi-objective design and decision-making process in water distribution systems. *Engineering Optimization*, Submitted.
- Batchabani, E., & Fuamba, M. (2014a). Multi-objective stochastic design of water distribution systems including water quality. *Journal of water resources planning and management*, Submitted.
- Batchabani, E., & Fuamba, M. (2014b). Optimal Tank Design in Water Distribution Networks: Review of Literature and Perspectives. *Journal of water resources planning and management*, 140(2), 136-145.
- Batchabani, E., Jalliffier-Verne, I., & Fuamba, M. (2012). Discussion of "Theory-Based SCS-CN Method and Its Applications" by W. H. Chung, I. T. Wang, and R. Y. Wang. *Journal of Hydrologic Engineering*, 17(2), 354-355.
- Belegundu, A., Murthy, D., Salagame, R., & Constans, E. (1994). *Multi-objective optimization of laminated ceramic composites using genetic algorithms*.
- Benjamin, J. R., & Cornell, C. A. (1970). *Probability, statistics, and decision for civil engineers*: McGraw-Hill, New York.

- Bolognesi, A., Bragalli, C., Marchi, A., & Artina, S. (2010). Genetic Heritage Evolution by Stochastic Transmission in the optimal design of water distribution networks. *Advances in Engineering Software*, 41(5), 792-801.
- Brière, F. (2007). *Drinking-Water Distribution, Sewage, and Rainfall Collection*. Montréal, Québec: Presses inter Polytechnique. 395p.
- Broad, D. R., Maier, H. R., Dandy, G. C., & Nixon, J. B. (2006). *Optimal Design of Water Distribution Systems Including Water Quality and System Uncertainty*. Paper presented at the Water Distribution Systems Analysis Symposium 2006.
- Chevalier, J. (2002). *Influence du réservoir McTavish sur la qualité de l'eau potable dans le réseau de distribution de la ville de Montréal*.
- Cohon, J. L. (2004). *Multiobjective programming and planning*. New York: Dover Publications.
- da Conceicao Cunha, M., & Sousa, J. (1999). Water distribution network design optimization: simulated annealing approach. *Journal of water resources planning and management*, 125(4), 215-221.
- Damelin, E., Shamir, U., & Arad, N. (1972). Engineering and economic evaluation of the reliability of water supply. *Water Resources Research*, 8(4), 861-877.
- Dandy, G., Simpson, A., & Murphy, L. (1996). An improved genetic algorithm for pipe network optimization. *Water Resources Research*, 32(2), 449-458.
- De Neufville, R., Schaake, J., & Stafford, J. H. (1971). Systems analysis of water distribution networks. *Journal of the Sanitary Engineering Division*, 97(6), 825-842.
- Djebedjian, B., Herrick, A., & Rayan, M. A. (2000). "An Investigation of the Optimization of Potable Water Network". Paper presented at the 5th International Water Technology Conference IWTC 2000, Alexandria, Egypt.
- Doby, T. A., Kumar, S. V., Baugh, J. W., Brill, E. D., & Ranjithan, S. R. (2004). *Genetic algorithm search for least cost design of looped pipe networks using age as a quality surrogate and different levels of redundancy*. Paper presented at the World Water and Environmental Resources Congress 2001.
- Duan, N., Mays, L., & Lansey, K. (1990). Optimal Reliability-Based Design of Pumping and Distribution Systems. *Journal of Hydraulic Engineering*, 116(2), 249-268.
- Ekinci, Ö., & Konak, H. (2009). An Optimization Strategy for Water Distribution Networks. *Water Resources Management*, 23(1), 169-185.
- Eusuff, M. M., & Lansey, K. E. (2003). Optimization of water distribution network design using the shuffled frog leaping algorithm. *Journal of water resources planning and management*, 129(3), 210-225.
- Fang, H., Zhang, J., & Gao, J. (2010). Optimal operation of multi-storage tank multi-source system based on storage policy. *Journal of Zhejiang University-Science A*, 11(8), 571-579.
- Farmani, R., Savic, D. A., & Walters, G. A. (2004). *The simultaneous multi-objective optimization of anytown pipe rehabilitation, tank sizing, tank siting and pump operation schedules*. Paper presented at the 2004 World Water and Environmental Resources

- Congress: Critical Transitions in Water and Environmental Resources Management, June 27, 2004 - July 1, 2004.
- Farmani, R., Walters, G., & Savic, D. (2006). Evolutionary multi-objective optimization of the design and operation of water distribution network: total cost vs. reliability vs. water quality. *Journal of Hydroinformatics*, 8(3), 165-179.
- Farmani, R., Walters, G. A., & Savic, D. A. (2005). Trade-off between total cost and reliability for Anytown water distribution network. *Journal of Water Resources Planning and Management-Asce*, 131(3), 161-171.
- Filion, Y., Adams, B., & Karney, B. (2007a). Cross correlation of demands in water distribution network design. *Journal of water resources planning and management*, 133(2), 137-144.
- Filion, Y. R. (2008). Single-objective deterministic versus multi-objective stochastic water network design: practical considerations for the water industry. *Proceedings of the Institution of Mechanical Engineers, Part O (Journal of Risk and Reliability)*, 222(Institution of Engineering and Technology, U.K.), 655-665.
- Filion, Y. R., Adams, B. J., & Karney, B. W. (2007b). Stochastic design of water distribution systems with expected annual damages. *Journal of water resources planning and management*, 133(3), 244-252.
- Fire-Underwriters-Survey. (1999). *Water supply for public fire protection*. Retrieved April 27, 2014. from <http://www.scm-rms.ca/docs/Fire%20Underwriters%20Survey%20-%201999%20Water%20Supply%20for%20Public%20Fire%20Protection.pdf>.
- Fonseca, C. M., & Fleming, P. J. (1993). *Genetic algorithms for multiobjective optimization: formulation, discussion and generalization*. Paper presented at the Proceedings of ICGA-93: Fifth International Conference on Genetic Algorithms, 17-22 July 1993, San Mateo, CA, USA.
- Fujiwara, O. (1990). A two-phase decomposition method for optimal design of looped water distribution networks. *Water Resources Research*, 26(4), 539-549.
- Gauthier, V., Besner, M., Barbeau, B., Millette, R., & Prévost, M. (2000). Storage tank management to improve drinking water quality: Case study. *Journal of water resources planning and management*, 126(4), 221-228.
- Geem, Z. (2006). Optimal cost design of water distribution networks using harmony search. *Engineering Optimization*, 38(3), 259-277.
- Ghimire, S. R., & Barkdoll, B. D. (2010). Sensitivity analysis of municipal drinking water distribution system energy use to system properties. *Urban Water Journal*, 7(4), 217 - 232.
- Giustolisi, O., Laucelli, D., & Colombo, A. F. (2009). Deterministic versus stochastic design of water distribution networks. *Journal of water resources planning and management*, 135(2), 117-127.
- Goldberg, D. E. (1987). Computer-aided pipeline operation using genetic algorithms and rule learning. Part I: Genetic algorithms in pipeline optimization. *Engineering with Computers*, 3(1), 35-45.

- Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*: Addison-wesley.
- Hodnett, R. M. (1981). Water Storage facilities and suction supplies. *Fire Protection Handbook*, National Fire Protection Association, Quincy, MA.
- Homwongs, C., Sastri, T., & Foster Iii, J. W. (1994). Adaptive forecasting of hourly municipal water consumption. *Journal of water resources planning and management*, 120(6), 888-904.
- Hydro-Quebec. (2013). Rates for business customers. Retrieved February 04, 2014, from <http://www.hydroquebec.com/business/rates-and-billing/rates/electricity-rates-business-customers/rate-m/>
- Iman, R. L., & Conover, W. J. (1982). A distribution-free approach to inducing rank correlation among input variables. *Communications in Statistics-Simulation and Computation*, 11(3), 311-334.
- InfraGuide. (2005). *Gestion d'un actif d'infrastructure - Outil d'apprentissage*. Canada.
- IWSA. (1955). *Subject no.3. Design and construction of service reservoirs and water towers*. Paper presented at the Proceedings of the Third Congress.
- Jung, D., Chung, G., & Kim, H. (2010, Sept. 12-15). *Optimal design of water distribution systems considering uncertainties in demands and roughness coefficients*. Paper presented at the Proceedings of the 12th Annual Water Distribution Systems Analysis Conference WDSA2010, Tucson, AZ, USA.
- Kapelan, Z. S., Savic, D. A., & Walters, G. A. (2005). Multiobjective design of water distribution systems under uncertainty. *Water Resources Research*, 41(11), W11407.
- Khu, S.-T., & Keedwell, E. (2005). Introducing more choices (flexibility) in the upgrading of water distribution networks: The New York city tunnel network example. *Engineering Optimization*, 37(3), 291-305.
- Klempous, R., Kotowski, J., Nikodem, J., & Ulasiwicz, J. (1997). Optimization algorithms of operative control in water distribution systems. *Journal of Computational and Applied Mathematics*, 84(1), 81-99.
- Kretzmann, H. A., & Van Zyl, J. E. (2004). *Stochastic analysis of water distribution systems*. Paper presented at the 2004 World Water and Environmental Resources Congress: Critical Transitions in Water and Environmental Resources Management, June 27, 2004 - July 1, 2004, Salt Lake City, UT, United states.
- Lansey, K., & Mays, L. (1989). Optimization model for water distribution system design= Modèle d'optimisation pour la conception de systèmes de distribution de l'eau. *Journal of Hydraulic Engineering*, 115(10), 1401-1418.
- Lansey, K. E., Duan, N., Mays, L. W., & Tung, Y. K. (1989b). Water distribution system design under uncertainties. *Journal of water resources planning and management*, 115(5), 630-645.
- Liang, T. (1971). Design conduit system by dynamic programming. *Journal of the Hydraulics Division*, 97(3), 383-393.

- Maier, H., Simpson, A., Zecchin, A., Foong, W., Phang, K., Seah, H., et al. (2003). Ant colony optimization for design of water distribution systems. *Journal of water resources planning and management*, 129(3), 200-209.
- Mallor, F., Mateo Collazos, P. M., Alberto Moralejo, I., & Azcarate Gimenez, C. (2003). *Multiobjective evolutionary algorithms: Pareto rankings*.
- Marchi, A., Dandy, G., Wilkins, A., & Rohrlach, H. (2014). A Methodology for Comparing Evolutionary Algorithms for the Optimization of Water Distribution Systems. *Journal of water resources planning and management*, 140(1), 22-31.
- Marchi, A., Salomons, E., Ostfeld, A., Kapelan, Z., Simpson, A. R., Zecchin, A. C., et al. (2013). The Battle of the Water Networks II (BWN-II). *Journal of water resources planning and management*, 10.1061/(ASCE)WR.1943-5452.0000378(May. 18, 2013).
- Mays, L. W. (2000). *Water distribution systems handbook*: McGraw-Hill.
- McKay, M. D., Beckman, R. J., & Conover, W. J. (2000). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 42(1), 55-61.
- McPherson, M. B. (1966). Ground storage booster pumping hydraulics. *American Water Works Association -- Journal*, 58(3), 311-325.
- MDDEP. (2002). *Directive 001 : captage et distribution de l'eau*. Retrieved April 27, 2014. from <http://www.mddefp.gouv.qc.ca/eau/potable/installation/documents/Directive001.pdf>.
- MDDEP. (2006). *Guide de conception des installations de production d'eau potable (Volume 1)*. Retrieved February 04, 2014. from <http://www.mddep.gouv.qc.ca/eau/potable/guide/documents/volume1.pdf>.
- Morley, M. S., & Tricarico, C. (2008). *Pressure Driven Demand Extension for EPANET (EPANETpdd)*: Technical Report 2008-02, University of Exeter, Centre for Water Systems.
- Murphy, L. J., Dandy, G. C., & Simpson, A. R. (1994). *Optimum Design and Operation of Pumped Water Distribution-Systems*. Paper presented at the 1994 International Conference on Hydraulics in Civil Engineering - Hydraulics Working with the Environment.
- Nel, D., & Haarhoff, J. (1996). Sizing municipal water storage tanks with Monte Carlo simulation. *Journal of Water Supply Research and Technology-Aqua*, 45(4), 203-212.
- Oliveira Sousa, J. J., Conceicao Cunha, M., & Almeida Sa Marques, J. A. (2005). *Simulated annealing reaches Anytown*. Paper presented at the Proc., 8th Int. Conf. CCWI05, Water Management for the 21st Century.
- Ostfeld, A. (2005). Optimal design and operation of multiquality networks under unsteady conditions. *Journal of water resources planning and management*, 131(2), 116-124.
- Pagé, R. (2005). *Plan directeur du réseau de distribution d'eau de la ville de Saint-Jérôme: Sainte-Foy*: Groupe Tanguay, Montréal, Canada.
- Piqueiro, F., Tentugal-Valente, J., & Tomas, M. (1991). Considerations about the storage capacity in water supply systems(influence of network consumption law and pumping

- station operation in its optimization). *Computational Mechanics Publications, Southampton, U.K.*, 171-183.
- Prasad, T. D. (2004). Multiobjective genetic algorithms for design of water distribution networks. *Journal of water resources planning and management*, 130(1), 73-82.
- Prasad, T. D. (2007). *Design of anytown network with improved tank sizing methodology*. Paper presented at the Combined International Conference of Computing and Control for the Water Industry CCWI2007 and Sustainable Urban Water Management SUWM2007, September 3, 2007 - September 5, 2007.
- Prasad, T. D. (2010). Design of Pumped Water Distribution Networks with Storage. *Journal of Water Resources Planning and Management-Asce*, 136(1), 129-132.
- Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. (1990). *Numerical recipes: the art of scientific computing* New York: Cambridge university press.
- Quindry, G., Brill Jr, E., Liebman, J., & Robinson, A. (1979). Comment on Design of Optimal Water Distribution Systems' by E. Alperovits and U. Shamir. *Water Resources Research*, 15(6), 1651-1654.
- Quindry, G., Liebman, J., & Brill, E. (1981). Optimization of looped water distribution systems. *Journal of the Environmental Engineering Division*, 107(4), 665-679.
- Reca, J., & Martínez, J. (2006). Genetic algorithms for the design of looped irrigation water distribution networks. *Water Resour. Res.*, 42(5), W05416.
- Roshani, E., MacLeod, S. P., & Filion, Y. R. (2012). Evaluating the Impact of Climate Change Mitigation Strategies on the Optimal Design and Expansion of the Amherstview, Ontario, Water Network: Canadian Case Study. *Journal of water resources planning and management*, 138(2), 100-110.
- Rossman, L. A. (1999). Computer models/EPANET. *Water distribution systems handbook*, New York: McGraw Hill.
- Rossman, L. A. (2000). EPANET 2: users manual.
- Savic, D. (2002). *Single-objective vs. multiobjective optimisation for integrated decision support*. Paper presented at the Proc., First Biennial Meeting of the Int. Environmental Modelling and Software Society.
- Savic, D., Kapelan, Z., Farmani, R., & Giustolisi, O. (2008). Optimal design and management of water distribution systems. In *Numerical Modelling of Hydrodynamics for Water Resources: Proceedings of the Conference on Numerical Modelling of Hydrodynamic Systems* (pp. 25-45): Taylor and Francis/Balkema, Leiden, Netherlands.
- Savic, D., & Walters, G. (1997). Genetic algorithms for least-cost design of water distribution networks. *Journal of water resources planning and management*, 123(2), 67-77.
- Sherali, H. D., Totlani, R., & Loganathan, G. (1998). Enhanced lower bounds for the global optimization of water distribution networks. *Water Resources Research*, 34(7), 1831-1841.
- Shu-li, Y., Nien-Sheng, H., Louie, P. W. F., & Yeh, W. W. G. (1996). Water distribution network reliability: connectivity analysis. *Journal of Infrastructure Systems*, 2(2), 54-64.

- Simpson, A. R., Dandy, G. C., & Murphy, L. J. (1994). Genetic algorithms compared to other techniques for pipe optimization. *Journal of water resources planning and management*, 120(4), 423-443.
- Todini, E. (2000). Looped water distribution networks design using a resilience index based heuristic approach. *Urban water*, 2(2), 115-122.
- Tolson, B., Maier, H., Simpson, A., & Lence, B. (2004). Genetic Algorithms for Reliability-Based Optimization of Water Distribution Systems. *Journal of water resources planning and management*, 130(1), 63-72.
- Tucciarelli, T., Criminisi, A., & Termini, D. (1999). Leak analysis in pipeline systems by means of optimal valve regulation. *Journal of Hydraulic Engineering*, 125(3), 277-285.
- Vamvakeridou-Lyroudia, L. S., Savic, D. A., & Walters, G. A. (2007). Tank simulation for the optimization of water distribution networks. *Journal of Hydraulic Engineering-Asce*, 133(6), 625-636.
- Vamvakeridou-Lyroudia, L. S., Walters, G. A., & Savic, D. A. (2005). Fuzzy Multiobjective Optimization of Water Distribution Networks. *Journal of water resources planning and management*, 131(6), 467-476.
- van Zyl, J. E., & Haarhoff, J. (1997). South African fire water guidelines and their impact on water supply system cost. *J. South African Institution of Civil Engineering*, 39(1), 49-53.
- van Zyl, J. E., Legat, Y., Piller, O., & Walski, T. M. (2012). Impact of Water Demand Parameters on the Reliability of Municipal Storage Tanks. *Journal of Water Resources Planning and Management*, 138(5), 553-561.
- van Zyl, J. E., Piller, O., & le Gat, Y. (2008a). Sizing municipal storage tanks based on reliability criteria. *Journal of water resources planning and management*, 134(6), 548-555.
- van Zyl, J. E., Piller, O., & Legat, Y. (2008b). *Critical parameters for the reliability of municipal storage tanks*. Paper presented at the World Environmental and Water Resources Congress 2008: Ahupua'a, May 12, 2008 - May 16, 2008.
- Wagner, J. M., Shamir, U., & Marks, D. H. (1988). Water distribution reliability: simulation methods. *Journal of water resources planning and management*, 114(3), 276-294.
- Wall, M. (1996). GALib: A C++ library of genetic algorithm components. *Mechanical Engineering Department, Massachusetts Institute of Technology*, 87.
- Walski, T. (2001). The wrong paradigm-Why water distribution optimization doesn't work. *Journal of water resources planning and management*, 127(4), 203-205.
- Walski, T., Youshock, M., & Rhee, H. (2000). *Use of modeling in decision making for water distribution master planning*. Paper presented at the Proceedings of the ASCE EWRI Conference, Minneapolis, Minnesota.
- Walski, T. M. (1995). Optimization and Pipe-Sizing Decisions. *Journal of Water Resources Planning and Management-Asce*, 121(4), 340-343.
- Walski, T. M. (2000). Hydraulic design of water distribution storage tanks. *Water distribution systems handbook*, L. Mays, ed, 10.11-10.20.

- Walski, T. M., Brill, J. E. D., Gessler, J., Goulter, I. C., Jeppson, R. M., Lansey, K., et al. (1987). Battle of the Network Models: Epilogue. *Journal of water resources planning and management*, 113(2), 191-203.
- Walters, G. A., Halhal, D., Savic, D., & Ouazar, D. (1999). Improved design of "Anytown" distribution network using structured messy genetic algorithms. *Urban water*, 1(1), 23-28.
- Wood, D. J. (1980). *Computer Analysis of Flow in Pipe Networks Including Extended Period Simulations: User's Manual*: Office of Continuing Education and Extension of the College of Engineering of the University of Kentucky.
- Wu, W., Maier, H. R., & Simpson, A. R. (2010b). Single-objective versus multiobjective optimization of water distribution systems accounting for greenhouse gas emissions by carbon pricing. *Journal of water resources planning and management*, 136(5), 555-565.
- Wu, W., Simpson, A. R., & Maier, H. R. (2010a). Accounting for greenhouse gas emissions in multiobjective genetic algorithm optimization of water distribution systems. *Journal of water resources planning and management*, 136(2), 146-155.
- Wu, Z. Y., Walski, T. M., Naumick, G., Dugandzic, J., & Nasuti, R. (2005). *Optimizing water system improvement for a growing community*. Paper presented at the International Conference of Computing and Control in the Water Industry, September 5-7.

ANNEXE 1 – Format du fichier d'entrée EPANET

```

[TITLE]
ANYTOWN example

[JUNCTIONS]
;ID Noeud      Altitude      Demande Base      Courbe Modulation
1              20              500              1
2              50              200              1
3              50              200              1
4              50              600              1
5              80              600              1
6              80              600              1
7              80              600              1
8              80              400              1
9              120             400              1
10             120             400              1
11             120             400              1
12             50              500              1
13             50              500              1
14             50              500              1
15             50              500              1
16             120             400              1
17             120             1000             1
18             50              500              1
19             50              1000             1
20             20              0                1
21             50              0                1
22             120             0                1

[RESERVOIRS]
;ID Noeud      Charge Tot      Courbe Modulation
40             10
;

[TANKS]
;ID Noeud      Altitude      niv Init      niv Min      niv Max      Diamètre      volMin      Courbe vol
41             215             0              0              35           32.65         0
42             215             0              0              35           32.65         0
43             250             0              0              35           35            0

[PIPES]
;ID Arc      Noeud1      Noeud2      Longueur      Diamètre      Rugosité      PerSing      État
1            1          2          12000         12           120          0          Open
2            1          12         12000         30           130          0          Open
3            1          13         12000         16           70           0          Open
4            1          20          100           30           130          0          Open
5            2          3           6000         10           120          0          Open
6            2          4           9000         10           120          0          Open
7            2          13          9000         12           70           0          Open
8            2          14          6000         10           120          0          Open
9            3          4           6000         10           120          0          Open
10           4          5           6000         12           130          0          Open
11           4          8          12000         8            120          0          Open
12           4          15          6000         10           120          0          Open
13           5          6           6000         6            130          0          Open

14           6          7           6000         6            130          0          Open
15           6          8           6000         12           130          0          Open
16           7          8           6000         10           130          0          Open
17           8          9          12000         8            120          0          Open
18           8          15          6000         10           120          0          Open
19           8          16          6000         8            120          0          Open
20           8          17          6000         18           120          0          Open
21           9          10          6000         8            120          0          Open
22           10         11          6000         8            120          0          Open
23           10         17          6000         10           120          0          Open
24           11         12          6000         8            120          0          Open
25           11         17          6000         6            130          0          Open
26           12         17          6000         20           120          0          Open
27           12         18          6000         10           130          0          Open
28           13         14          6000         12           70           0          Open
29           13         18          6000         12           70           0          Open
30           13         19          6000         10           70           0          Open
31           14         15          6000         12           70           0          Open
32           14         19          6000         10           70           0          Open
33           14         21          100          12           120          0          Open
34           15         16          6000         10           70           0          Open
35           15         19          6000         10           70           0          Open
36           16         17          6000         8            120          0          Open
37           16         18          6000         12           70           0          Open
38           16         19          6000         10           70           0          Open
39           17         18          6000         8            120          0          Open
40           17         22          100          12           120          0          Open
41           18         19          6000         10           70           0          Open
42           21         41          1            12           120          0          Open
43           22         42          1            12           120          0          Open
44           43         8            1            12           100          0          Open

[PUMPS]
;ID Arc      NoeudAsp      NoeudDéch      Paramètres
78           40           20             HEAD 2 PATTERN 2
79           40           20             HEAD 2 PATTERN 3
80           40           20             HEAD 2 PATTERN 4

[VALVES]
;ID Arc      NoeudAmon      NoeudAval      Diamètre      Type      Consigne      PerSing

[TAGS]

[DEMANDS]
;ID Noeud      Demande Base      Courbe Modulation      Catégorie

[STATUS]
;ID Arc      État/Consigne

```

```

[PATTERNS]
;ID Modulation      Multiplicateurs
;
1      1      1      1      0.9      0.9      0.9
1      0.7    0.7    0.7    0.6      0.6      0.6
1      1.2    1.2    1.2    1.3      1.3      1.3
1      1.2    1.2    1.2    1.1      1.1      1.1
;
2      1      1      1      1      1      1
2      1      1      1      1      1      1
2      1      1      1      1      1      1
2      1      1      1      1      1      1
;
3      1      1      1      1      1      1
3      1      1      1      1      1      1
3      1      1      1      1      1      1
3      1      1      1      1      1      1
;
4      0      0      0      0      0      0
4      0      0      0      0      0      0
4      1      1      1      1      1      1
4      1      1      1      1      1      0
;
5      1      1      1      1      1      1
5      1      1      1      1      1      0
5      1      1      1      1      1      1
5      1      1      1      1      1      1

[CURVES]
;ID Courbe      valeur X      valeur Y
;PUMP:          ;
2      0      300
2      2000   292
2      4000   270
2      6000   230
2      8000   181
;EFFICIENCY:    ;
1      0      0
1      2000   50
1      4000   65
1      6000   55
1      8000   40

[CONTROLS]

[RULES]
[ENERGY]
Global Efficiency 75
Global Price      0.12
Pump 78 Efficiency 1
Pump 79 Efficiency 1
Pump 80 Efficiency 1
Demand Charge     0

[EMITTERS]
;ID Noeud      Coefficient

[QUALITY]
;ID Noeud      Qualité Initiale

[SOURCES]
;ID Noeud      Type      Qualité      Courbe Modul

[REACTIONS]
;Type      Tuyau/Réservoir      Coefficient
Tank      42      0
Tank      43      0

[REACTIONS]
Order Bulk      1
Order Tank      1
Order wall      1
Global Bulk     0
Global wall     0
Limiting Potential 0
Roughness Correlation 0

[MIXING]
;ID Réservoir      Modèle      Fraction Mélange

[TIMES]
Duration      24:00:00
Hydraulic timestep 00:02
Quality timestep 00:01
Pattern timestep 01:00
Pattern start 00:00
Report timestep 01:00
Report start 00:00
Start ClockTime 12:00 AM
Statistic      None

```

[REPORT]
 Status No
 Summary Yes
 Page 0
 Energy Yes
 Nodes All
 Links All
 State Yes

[OPTIONS]
 Units GPM
 Headloss H-W
 Specific Gravity 1
 Viscosity 1
 Trials 40
 Accuracy 0.001
 Unbalanced Continue 10
 Pattern 1
 Demand Multiplier 1
 Emitter Exponent 0.5
 Quality Age
 Diffusivity 1
 Tolerance 0.01

[COORDINATES]
 ;ID Noeud Coord X Coord Y
 1 7682.33 3371.15
 2 7633.71 5737.44
 3 7520.26 7293.35
 4 6175.04 7568.88
 5 5676.79 8454.66
 6 4100.49 8071.31
 7 3192.87 7763.37
 8 3321.19 6696.55
 9 1305.62 4850.97
 10 2317.45 3588.2
 11 3701.63 2892.06
 12 4846.03 3354.94
 13 6450.57 4424.64
 14 6466.77 5769.85
 15 5094.45 6482.09
 16 4017.86 5615.96
 17 3005.49 4292.43
 18 4846.03 4440.84
 19 5332.25 5332.25
 20 8544.18 3371.15
 21 6094 6499.19
 22 2422.68 4519.09
 40 8700.6 3371.15
 41 5488.01 6543.89
 42 2292.77 4114.29

43 2857.95 6883.05

[VERTICES]
 ;ID Arc Coord X Coord Y

[LABELS]
 ;Coord X Coord Y Texte et Noeud d'Ancre

[BACKDROP]
 DIMENSIONS 935.87 2613.65 9070.35 8738.7
 UNITS None
 FILE
 OFFSET 0 0

[END]

ANNEXE 2 – Description complète du problème d'optimisation du réseau de D-Town

4. Problem description

The municipality of D-Town is in need of a design project to cope with the increased water demand of the population and a general increase in the population that has led to a new residential district in the Eastern part of the city (junctions from N1 to N10 in the EPANET file, where the demand pattern assigned is DMA3_pat). To accomplish this task the city has already commissioned the development of a calibrated hydraulic model of the actual network so as to better evaluate its present state and its future behaviour (as outlined above, this file can be downloaded from <http://wdsa2012.com/>). The network model includes the forecasted demands, the demand patterns and the layout of the new district. It also contains existing pump and tank characteristics and the actual controls of pumps and valves. The model shows that the existing infrastructure is not able to meet the forecasted demand and therefore an upgrade of the network is necessary.

The water utility is interested in minimizing operational and capital costs, but is also interested in reducing the greenhouse emissions, as it is likely that cap and trade policy or a carbon tax will be introduced in the future. Moreover, as the actual network model shows that the water is in the network for long periods of time before reaching the users, the utility is also interested in reducing the water age in the network.

The specific criteria required by the water utility to assess the project design and the design options are outlined below.

4.1 Design Requirements

4.1.1 Costs

The water utility desires a low operational and capital cost design. In particular, the operational costs are a result of pump operations only, while capital costs are associated with the pipe and tank material and construction, and any upgrades of the existing pumping stations.

As capital and operational costs occur at different time points during the lifetime of the project, annual costs are provided for the new possible components of the network in the section *Design Options*. These costs take into account the lifetime of each specific component and the discount rate. Consequently, the total cost requiring minimization is the sum of the annual capital costs and of the annual cost of pumping operations.

To account for the annual cost of pumping operations, the weekly cost of the pump power has to be computed. This cost is then multiplied by the number of weeks in a year (52) and, to account for the demand variability throughout the year, this number is divided by the peak-day factor of 1.3. (The variability of the electric tariff, of the demand, and of any other design variable, during the year and the lifetime of the project is not considered.)

4.1.2 Water age

The water age is considered by the utility to be too high in some parts of the network, and the utility desires that the water age within the project design to be kept to a minimum. The utility has defined the measure of network water age as follows:

$$WA_{net} = \frac{\sum_{i=1}^{N_{junc}} \sum_{j=1}^{N_{time}} k_{ij} Q_{dem,ij} WA_{ij}}{\sum_{i=1}^{N_{junc}} \sum_{j=1}^{N_{time}} Q_{dem,ij}}$$

where WA_{net} is the network water age (hours), WA_{ij} is the water age (hours) at junction i (note: tanks/reservoirs excluded) at time t_j , k_{ij} is the variable defined as follows:

$$k_{ij} = \begin{cases} 1 & , \quad WA_{ij} \geq WA_{th} \\ 0 & , \quad WA_{ij} < WA_{th} \end{cases}$$

where WA_{th} is the water age threshold equal to 48 hours (for this competition), $Q_{dem,ij}$ is the demand at junction i and time t_j , i is the junction index ($i=1, 2, \dots, N_{junc}$), j is the time index ($j=0,1,\dots,N_{time}$), t_j is the simulation time $t_j=j\Delta t$, Δt is the time step equal to 1 hour (resulting in all water age and demand variables to be computed only on the hour), N_{junc} is the number of system junctions and N_{time} is the number of simulation time steps (equal to 168 as the extended period simulation time is one week). WA_{net} should be assessed during normal system operation only.

Note that the above network water age measure considers water age above the threshold only which is compatible with AWWA/EPA guidelines. It also considers water age at non-zero demand nodes only and gives more weight to nodes with larger demands (implying a risk type indicator). Finally, the above network water age measure takes into account water age at junctions only (the age of water in a tank is taken into account indirectly, via corresponding demand junctions supplied from that tank).

4.1.3 Greenhouse gas emissions

The water utility foresees that the introduction of a carbon tax, or a cap and trade policy, is likely to occur in the near future. Hence they would like to minimize the greenhouse gas (GHG) emissions associated with the design.

From the water utility's perspective, GHGs can be subdivided into (a) capital GHG emissions (associated with the energy used to produce, transport and install pipes, tanks and pumps), and (b) operational GHG emissions (resulting from the operation of pumps drawing their electricity from fossil fuel sources). In particular, the water utility wants to minimize the annualized total GHG emissions, which are considered as the sum of the annualised capital GHG emissions, and of the annual operational GHG emissions.

The capital component of the annual GHG emission is estimated using Table 2, which represents the annualised CO₂-equivalent emissions to manufacture a meter of a pipe.

Table 2. Annual CO₂ equivalent emissions to manufacture the pipes.

Diameter (mm)	Annualised EE (kg-CO ₂ -e/m/year)
102	5.90
152	9.71
203	13.94
254	18.43
305	23.16
356	28.09
406	33.09
457	38.35
508	43.76
610	54.99
711	66.57
762	72.58

Annual operational GHG emissions are computed as the product of total annual pump energy consumption in the system under normal operation (see section 4.1.1) and the emission factor of 1.04 kg-CO₂-e/kWh.

Additional information:

As in the case of annual capital costs, values shown in Table 2 already take into account the lifetime of the components and a discount factor. The discount factor here has been chosen to be 0% as recommended by the Intergovernmental Panel in Climate Change (IPCC). Emissions to manufacture pumps and tanks will be neglected due to unavailability of accurate data at this stage.

The emission factor to convert the energy into a carbon equivalent (1.04 kg-CO₂-e/kWh) is assumed to be constant. Although the emission factor changes according to the sources used to produce the energy (and therefore according to the time of the day, the period of the year and region, and it will also change in the future as the mix of renewable energy sources changes), a constant value is considered for simplicity.

4.1.4 System performance under normal operation

The water utility requires that every demand node of the network has water delivered to it with adequate pressure. Nodes without demand only have the requirement of a minimum pressure above zero. The minimum pressure required for nodes with demands is 25 m.

An additional requirement of the water utility is that at the end of the extended period simulation (1 week) each tank has to have at least the same volume of water it had at the beginning of the simulation. Note that this initial volume has to be set equal to half the volume of the tank (see Figure 1). Moreover, during normal operation, tanks are not allowed to stay empty.

In the EPANET2 input file, tanks are described by 6 values: diameter, elevation, minimum level, maximum level and initial level as shown in Figure 1. In EPANET2 the tank level is forced to stay between the minimum and the maximum tank level, so that the volume between the tank bottom elevation and the minimum level represents a reserve volume. If a simulation requires water from a tank that is already at its minimum level, EPANET2 disconnects the tank so that the tank level will remain at its minimum value.

The requirement of tanks not staying empty can be therefore translated in EPANET2 by checking that the tank water level (shown in the field “Pressure” in the tank property) is not equal to the minimum level for two or more consecutive time steps. In the input file provided, all tanks have a minimum level set equal to 0 m and hence the whole volume of the tank can be used. The tank minimum level can be changed as desired by the participants.

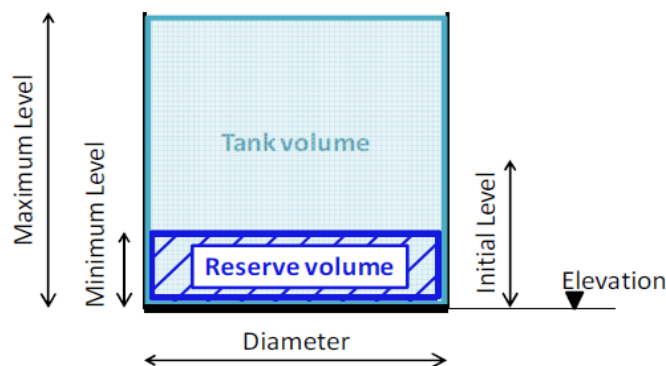


Figure 1. Tank description in EPANET2.

4.1.5 System performance under power failure operation

The provision of a reliable water supply is critical for the proposed network design. The water utility's major concern about failing to provide adequate supply is exclusively due to the possibility of a power outage. This is because a power outage will result in a shutdown of the network's pumps (except the pumps powered by the backup diesel generators), leading to an emptying of the supply tanks and the inadequate provision of pressure at the network nodes.

The water utility has characterised the design power failure event as being a two hour power outage. In the event of a power outage, the system is still expected to provide a minimum pressure of 25 m to nodes with a non-zero water demand; otherwise the minimum pressure required is zero. Moreover, tanks are not allowed to stay empty. Note that during power failure, tanks are allowed to operate below their minimum level if this value has been set larger than zero (that is, the reserve volume can be used during the power failure operation).

The power outage can occur at any time during the simulation week. Therefore, for a power failure scenario starting at time t , initial tank levels will be those determined from the normal operation conditions at time t , and the demands for the two hour duration of the power failure will be consistent with those specified in the EPANET2 input file provided for this time duration.

Please note that, in case of a power failure, the network performance will be assessed only for the two hours of power outage. Note also that the network performance can be improved by increasing the network storage volume and/or by introducing diesel generators (see Sections 4.2.2, and 4.2.4). If one or more diesel generators are selected, the pumps connected to them will be assumed constantly switched "ON" for two hours, i.e. for the full duration of power outage. Clearly, if a pump is not connected to a power generator, it will be switched "OFF" for those two hours.

4.2 Design Options

4.2.1 Pipes

Pipe diameter options and costs for the network expansion are given in Table 3. The costs shown are inclusive of pipe construction, transport and installation.

According to the water utility, pipes can also be placed in parallel to existing pipes (pipe IDs starting with the letter P in the EPANET2 input file). However, as this implies the disruption and reconstruction of pavement roads, the cost of duplicating existing pipes is given by the costs within Table, but with an additional cost premium of 20%.

Another requirement of the water utility is that none of the existing pipes can be decommissioned, with the exception of pipes with IDs equal to 1 and 2 in the EPANET2 input file. In fact, the water utility is not sure if it is more effective to link the new zone to the DMA3 or to the DMA2. Therefore one of the two pipes that link the new zone to the existing districts (pipes 1 and 2) can be omitted (a diameter option can be zero, meaning that the pipe is not built). The Hazen-Williams coefficient for every diameter is equal to 120.

Table 3. Pipe annual costs.

Diameter	Annual Cost New Pipe	Annual Cost Parall Pipe
(mm)	(\$/m/yr)	(\$/m/yr)
102	8.31	9.97
152	10.10	12.10
203	12.10	14.49
254	12.96	15.55
305	15.22	18.28
356	16.62	19.94
406	19.41	23.26
457	22.20	26.65
508	24.66	29.58
610	35.69	42.80
711	40.08	48.12
762	42.60	51.11

4.2.2 Tanks

Because of the increased demands, the water utility is also allowing for the addition of new tanks, but only adjacent to the existing tanks, where the water utility already owns sufficient land. New tanks are assumed to have the same height and bottom elevation as the existing adjacent tanks (because the water utility does not want to introduce new valves to control the system). All new tanks are cylindrical and come in pre-specified standard sizes shown in Table 4, together with associated annualised costs. The construction of non-standard tanks is discarded by the water utility because they are regarded as being too expensive.

Table 4. Tank annual costs.

Volume	Annual Cost
(m³)	(1000\$/yr)
500	14.02
1000	30.64
2000	61.21
3750	87.46
5000	122.42
10000	174.93

Note that the annual costs shown in Table 4 already include the connectivity costs to link the new tanks to the network. Therefore, the addition of new tanks can be modelled in EPANET simply by increasing the tank diameters so that the resulting volume is equal to the existing tank volume plus the new tank volume. Note also that all the constraints on operational levels of new and existing tanks have been already given in sections 4.1.4 and 4.1.5.

4.2.3 Pumps

Existing pump systems can be upgraded by replacing existing pumps with larger pumps or by adding new pumps to the existing pumping stations in parallel to the existing pumps. No additional pumping stations or boosters can be placed into the network as the water utility does not have any available location for these new components. Pump efficiency is assumed constant: existing pumps have an efficiency of 65%, while new pumps have an efficiency equal to 75%. Pump curves for available pumps are given in the Appendix in Tables A1-A10. Table 5 shows the annual costs associated with each pump.

Table 5. Pump annual costs.

Pump Model	Maximum Power	Annual Cost
	kW	(\$/yr)
8	45.24	4133
9	31.67	3563
10	49.76	4339
11	22.62	3225
8a	22.62	3225
10a	24.88	3307
11a	11.31	2850
8b	54.28	4554
9b	38.00	3820
10b	59.71	4823

4.2.4 Diesel generators

To improve the network's performances during power outage, the water utility allows for the purchasing and installation of diesel generators to be located within the existing pump stations. Note that more than one diesel generator can be inserted in the same pumping station. The costs given in Table 6 include generator costs, and related transport/installation costs. No costs associated with the land property needs to be considered.

Table 6. Diesel generator annual cost.

Power	Annual Cost
(kW)	(1000\$/yr)
50	9.45
100	10.56
200	11.63
300	15.00
400	16.78
800	25.74
1300	40.67
2500	62.29

Diesel generators should be installed to accommodate the total rated power of all pumps backed up by the generators, where the rated power of a pump corresponds to the maximum pump power given in Table 5. Energy consumption and GHG emissions associated with the use of diesel generators are considered negligible and hence should not be computed in the evaluation of performances related to the cost and the GHG emissions.

4.2.5 Valves

The water utility is not considering introducing any additional operational valves in the system. The pressure reducing valves currently present in the system are to be maintained (i.e. cannot be removed), but their operational settings can be modified if deemed necessary. However, their operational setting cannot be set above 60m. The only exception to the insertion of new valves is the possible installation of the PRV valve N15 in the new pipe 2 linking the new zone to the DMA3. If this valve is installed, its annualised cost should be determined from Table 7 for the pipe 2 diameter.

Table 7. PRV annual cost.

Diameter	Annual Cost
(mm)	(\$/yr)
102	323
152	529
203	779
254	1113
305	1892
356	2282
406	4063
457	4452
508	4564
610	5287
711	6122
762	6790

In all cases, modified settings of PRVs are assumed constant during simulation period, nor are they allowed to depend on the status of other network components). Note that the setting of existing PRVs and of the PRV N15 does not introduce any additional cost.

4.2.6 Pump controls

The existing pump controls shown in the EPANET input file provided can be changed without incurring additional costs. The same applies to changing the pump type control, i.e. switching from pumps controlled by the corresponding threshold tank levels to time controlled pumps or vice versa.

4.3 Electricity Tariff

The electricity tariff is shown in Table 8 where the energy prices are shown in cents/kWh. For example, the price applied on Monday is 6.72 cents/kWh from 6:00 am to 7:00 am, while it is 10.94 cents/kWh from 7:00 am to 8 am.

Table 8. Electricity tariff.

Hour/Day	Mon	Tue	Wed	Thu	Fri	Sat	Sun
0	6.72	6.72	6.72	6.72	6.72	6.72	6.72
1	6.72	6.72	6.72	6.72	6.72	6.72	6.72
2	6.72	6.72	6.72	6.72	6.72	6.72	6.72
3	6.72	6.72	6.72	6.72	6.72	6.72	6.72
4	6.72	6.72	6.72	6.72	6.72	6.72	6.72
5	6.72	6.72	6.72	6.72	6.72	6.72	6.72
6	6.72	6.72	6.72	6.72	6.72	6.72	6.72
7	10.94	10.94	10.94	10.94	10.94	10.94	6.72
8	10.94	10.94	10.94	10.94	10.94	10.94	6.72
9	10.94	10.94	10.94	10.94	10.94	10.94	6.72
10	27.68	27.68	27.68	27.68	27.68	10.94	6.72
11	27.68	27.68	27.68	27.68	27.68	10.94	6.72
12	27.68	27.68	27.68	27.68	27.68	10.94	6.72
13	27.68	27.68	27.68	27.68	27.68	10.94	6.72
14	27.68	27.68	27.68	27.68	27.68	10.94	6.72
15	27.68	27.68	27.68	27.68	27.68	10.94	6.72
16	27.68	27.68	27.68	27.68	27.68	10.94	6.72
17	10.94	10.94	10.94	10.94	10.94	10.94	10.94
18	10.94	10.94	10.94	10.94	10.94	10.94	10.94
19	10.94	10.94	10.94	10.94	10.94	10.94	10.94
20	10.94	10.94	10.94	10.94	10.94	6.72	10.94
21	6.72	6.72	6.72	6.72	6.72	6.72	6.72
22	6.72	6.72	6.72	6.72	6.72	6.72	6.72
23	6.72	6.72	6.72	6.72	6.72	6.72	6.72

5. Design evaluation

Each participant is required to submit one solution only - regardless of the methodology used. The solutions received will be ranked for each of the performance criteria of cost, water age, and GHG emissions and assigned a score depending on this ranking (given compliance to all performance criteria for both normal and power outage operations). The solution with the highest overall rank will be selected as the winner. Although they will be omitted from consideration, it is possible to describe other solutions that the participants have obtained in the paper.

Note that, as outlined earlier, to be eligible to participate, participants are required to submit: (i) a paper describing the approach adopted; (ii) the EPANET input file (*.inp) of the network solution selected (set up for the single week extended period simulation); and (iii) the completed Excel file outlining the design decisions made by the participants. Note also that the selected solution will be independently checked and evaluated using the hydraulic solver EPANET2 to verify the authors' results.

APPENDIX:

Available pump curves are shown in Tables A1-A10, where curves 8, 9, 10, 11 correspond to pumps currently operating in the network. It has to be noted that the pump curves cover the whole range of flows (from flow equal to 0 to the flow corresponding to zero head): the head-flow equation is also given in the tables where the equations are consistent with the three point curve fitting used by EPANET2.

Table A1. Pump curve 8:
 $H=70.00-0.07731 \cdot Q^{1.36}$.

Pump Model	Flow (L/s)	Head (m)
8	0	70
8	60	50
8	100	30

Table A2. Pump curve 9:
 $H=90.00-0.01331 \cdot Q^{2.15}$.

Pump Model	Flow (L/s)	Head (m)
9	0	90
9	30	70
9	50	30

Table A3. Pump curve 10:
 $H=120.00-0.001477 \cdot Q^{2.59}$.

Pump Model	Flow (L/s)	Head (m)
10	0	120
10	30	110
10	70	30

Table A4. Pump curve 11:
 $H=90.00-0.01104 \cdot Q^{2.41}$.

Pump Model	Flow (L/s)	Head (m)
11	0	90
11	30	50
11	40	10

Table A5. Pump curve 8a:
 $H=70.00-0.198 \cdot Q^{1.36}$.

Pump Model	Flow (L/s)	Head (m)
8a	0	70
8a	30	50
8a	50	30

Table A6. Pump curve 10a:
 $H=120.00-0.008915 \cdot Q^{2.59}$.

Pump Model	Flow (L/s)	Head (m)
10a	0	120
10a	15	110
10a	35	30

Table A7. Pump curve 11a:
 $H=90.00-0.05866 \cdot Q^{2.41}$.

Pump Model	Flow (L/s)	Head (m)
11a	0	90
11a	15	50
11a	20	10

Table A8. Pump curve 8b:
 $H=84.00-0.09277 \cdot Q^{1.36}$.

Pump Model	Flow (L/s)	Head (m)
8b	0	84
8b	60	60
8b	100	36

Table A9. Pump curve 9b:
 $H=108.00-0.01597 \cdot Q^{2.15}$.

Pump Model	Flow (L/s)	Head (m)
9b	0	108
9b	30	84
9b	50	36

Table A10. Pump curve 10b:
 $H=144.00-0.001773 \cdot Q^{2.59}$.

Pump Model	Flow (L/s)	Head (m)
10b	0	144
10b	30	132
10b	70	36

D-Town Network Layout

